DO INDIVIDUAL DIFFERENCES INTERACT WITH LEXICAL CUES DURING SPEECH RECOGNITION IN ADVERSE LISTENING CONDITIONS?

A thesis submitted in partial fulfilment of the

requirements for the Degree of

Master of Audiology

in the Department of Communication Disorders

at the University of Canterbury

by Sarah Elizabeth Kerr

2015

Table of Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I would like to thank my primary supervisor Associate Professor Megan McAuliffe for her support and feedback not only during the writing of this thesis, but also during my time as a research assistant in her lab. I would also like to thank my associate supervisor, Dr Don Sinex, who helped with the editing of the speech stimuli, and provided helpful comments on drafts. I would also like to thank my family, friends, classmates and fellow research assistants for their support.

Abstract

Purpose: This thesis examines the effect of listener characteristics (i.e., cognition and vocabulary) and language-based factors (i.e., lexical frequency and phonological similarity) on speech recognition accuracy in adverse listening conditions.

Method: Fifty listeners (40 females and 10 males) aged 18-33 years and with normal hearing (puretone thresholds ≤ 20 dB HL, 0.25-8 kHz) participated. They completed a speech perception experiment, which required listeners to repeat back non-sensical English phrases presented at a variety of signal-to-noise ratios (-5, -2, +1, and +4 dB SNRs). In addition, all listeners undertook assessments of vocabulary knowledge (PPVT-IV) and cognition (WAIS - IV). The primary dependent variable was individual content word recognition accuracy, and results were analysed using binomial mixed effects modelling.

Results: Listeners demonstrated variability in their speech recognition abilities, and their vocabulary and cognitive scores. Statistical analysis revealed that listener-based factors affected word recognition. Listeners with faster processing speed and larger working memories exhibited higher word recognition accuracy. Surprisingly, listeners with higher non-verbal intelligence scores exhibited lower word recognition accuracy. Vocabulary knowledge interacted with SNR, such that as the listening conditions became more favourable, listeners with larger receptive vocabularies identified more words correctly. Similarly, main effects were also present for language-based factors. The more phonologically distinct a word was, the more likely it was to be correctly identified; higher frequency words were more likely to be accurately recognised. In addition, higher frequency words were identified more accurately at higher SNR levels. Finally, listener- and language-based factors interacted. The positive effect of working memory on word recognition was reversed as word frequency increased; on the other

hand non-verbal intelligence's negative influence on word recognition was reversed as word frequency increased.

Conclusion: In the current cohort, listener and language-based factors interacted in the process of word recognition in noise. These results provide an insight into the underlying speech recognition mechanisms in adverse conditions. Further understanding of how these listener differences affect an individual's speech processing may lead to the development of improved signal processing techniques and rehabilitation strategies.

Introduction

Listening in Adverse Conditions

Everyday listening conditions are rarely optimal. In fact, Mattys, Davis, Bradlow and Scott (2012) note that speech recognition under adverse conditions is the rule rather than the exception. Thus, in order to understand speech recognition, it is important to determine the effects of adverse conditions on listeners' speech processing. According to Mattys et al. (2012) energetic masking has the most deleterious effects on listener comprehension. Energetic masking reduces intelligibility because the interfering signal (e.g., noise or other talkers) overlaps in frequency and time with the desired signal (e.g., talker of interest). This overlap renders portions of the desired signal inaudible (Brungart, 2001). Energetic masking is commonly experienced by listeners when, for example, they try to focus on a single talker in the presence of other talkers, or try to listen to the television over a portable fan.

Two of the most prominent effects of energetic masking on the listener are failure to recognise the acoustic-phonetic features of speech and perceptual interference (Mattys et al., 2012). Recognition failure arises when listeners mismatch acoustic-phonetic features to their correct segmental and therefore lexical representations. Recognition failure has various results: lexical uncertainty may arise due to more competition from lexically similar items; inaccurate lexical selection occurs; or the listener fails to choose a word at all. Perceptual interference, on the other hand, occurs when the desired signal competes with another incoming noise. This is similar to the effect of recognition failure; however, the competing noise also forces the listener to separate the signal from the noise in order to attend to the signal. Separating the signal from the noise requires listeners to have knowledge of both the signal and noise components of the input. Listeners must capitalise on changes in the signal to noise ratio of the input in order to glimpse the desired signal (Assmann & Summerfield, 2004). Once these glimpses have been obtained they require piecing together, which depends upon top down processes making use of context and redundant speech cues (Miller, Heise, & Lichten, 1951; Miller & Licklider, 1950). Thus, not only do the effects of recognition failure and perceptual interference pertain to cognitive demands placed on the listener attempting to decode the message, they also apply to linguistic information within the message. Both of these characteristics of speech perception will be explored further in the following sections. First the discussion will focus on the role of listener-specific characteristics, followed by the effects of linguistic cues on speech recognition in adverse conditions.

Individual Differences in Speech Recognition in Adverse Conditions

Listeners vary in their ability to recognise speech in adverse conditions (e.g., Gilbert, Tamati, & Pisoni, 2013; Mattys et al., 2012). To attempt to account for this listener variation, a number of studies have examined individual factors affecting performance in speech recognition—many of these focusing on the role of cognitive function in speech perception. The potential role of cognition was highlighted by the Working Group on Speech Understanding and Aging of the Committee on Hearing and Bioacoustics of Biomechanics (CHABA) report (1988), which suggested that older adults' difficulties in understanding speech in noise were partially accounted for by declines in their cognitive function—in particular, memory and processing speed. Since then working memory (e.g., Akeroyd, 2008; Desjardins & Doherty, 2013; Krull, Humes, & Kidd, 2013; McAuliffe, Gibson, Kerr, Anderson, & LaShell, 2013; Tamati, Gilbert, & Pisoni, 2013), processing speed (e.g., Neger, Rietveld, & Janse, 2014; Sommers & Danielson, 1999; Woods, Kalluri, Pentony, & Nooraei, 2013), and non-verbal intelligence (e.g., Conway, Bauernschmidt, Huang, & Pisoni, 2010; Tamati et al., 2013) have all been implicated in research attempting to understand the underlying cognitive mechanisms of speech recognition in adverse conditions. More recently, studies have begun to investigate whether a listener's linguistic skill (vocabulary knowledge) may also play a role in speech understanding in difficult listening situations (e.g., Benard, Mensink, & Baskent, 2014; McAuliffe et al., 2013; Tamati et al., 2013). Thus far, the findings for both cognition and vocabulary have been equivocal. These findings are discussed in turn below.

Working memory. Working memory refers to the ability to store and simultaneously process information (Baddeley, 1998; Daneman & Carpenter, 1980). Recent information is held in a mental workspace where this information is processed and integrated with long-term memory knowledge (Pichora-Fuller & Singh, 2006). During word recognition listeners must hold auditory input in working memory while they decide upon the best word match from possible lexical candidates in their long-term memory (Collison, Munson, & Carney, 2004). Thus, it is plausible to suggest that increased working memory capacity would aid word recognition in adverse listening conditions. However, the influence of working memory as a predictor of speech recognition ability appears to be variable—with some studies reporting a positive relationship between working memory and speech recognition (e.g., Desjardins & Doherty, 2013; Tamati et al., 2013) and others finding no link (e.g., Krull et al., 2013; McAuliffe et al., 2013). For example, working memory tasks, in particular reading span tasks (Daneman & Carpenter, 1980), have demonstrated that larger working memory is associated with better speech recognition abilities for both aided and unaided listeners in a number of studies (see review by Akeroyd, 2008). Younger normal hearing listeners' speech recognition was also predicted by their working memory abilities, as measured by backward digit span for highly variable sentence stimuli in noise (Tamati et al., 2013). Similarly, a study of younger and older listeners with and without hearing impairment (Desjardins & Doherty, 2013) also found that working memory, as measured by the reading span, was predictive of high and low context sentence in noise recognition.

However, other studies have failed to find effects of working memory on speech recognition in adverse conditions. Once again these studies vary widely in listener age, stimulus type, and hearing loss. For example, McAuliffe et al. (2013) found that working memory scores (digit span, letter number sequencing) were not associated with younger and older normal hearing listeners' perception of low context dysarthric speech. Similarly, Krull et al. (2013) reported no link between working memory and word recognition for unaided older listeners with near normal hearing (aged 65-84 years old, M = 73) nor aided older listeners with mild to moderately-severe sensorineural hearing loss (aged 60-85 years old, M = 75). Older adults, regardless of hearing ability, were poorer overall at speech recognition than the younger normal hearing adults in the study. These findings suggest that cognitive and sensory abilities, other than working memory and hearing loss, related to ageing are at play. However, it is difficult to draw further conclusions from Krull et al.'s (2013) study, since younger normal hearing listeners' working memory was not examined for comparison. Finally, Benard et al. (2014) found no relationship between working memory (as measured by the WAIS-IV (Wechsler, 2008)) and the identification of interrupted sentences by normal hearing listeners (aged 21 – 63 years).

From these studies the role of working memory on speech recognition is unclear. The studies discussed previously varied with respect to the age of listeners and listener hearing sensitivity. Yet other studies have found mixed results within participants—depending on stimulus type. For example, Ellis and Munro (2013) investigated the effects of frequency compressed low context sentences in noise for normal hearing listeners in relation to a working memory task. Overall, they found that reading span correlated with uncompressed speech in noise, but had no effect on listeners' ability to comprehend frequency compressed speech. The authors suggested that lack of effect between frequency compressed speech and working memory was related to the listeners' lack of experience with frequency compression.

It is also the case that listeners with normal hearing do not require frequency compression for speech sounds to be audible to them. Therefore, frequency compression could be providing too much of an unnatural speech stream for the normal hearing listeners to overcome. Ellis and Munro's (2013) findings suggest that it would be more beneficial to study the effects of cognition in environments which are more akin to everyday adverse listening conditions.

Processing speed. A listener's processing speed is also thought to be important for speech recognition in adverse conditions. Processing speed relates to the rate at which operations are successfully executed (Kaufman et al., 2010). Given the rate at which listeners must process speech, increased processing speed means that more information is available for higher level processing of the auditory signal (Schoof & Rosen, 2014). In general, studies have shown that when a listener has normal hearing, their processing speed bears no relationship to their ability to recognise speech in adverse conditions—for neither interrupted speech (Benard et al., 2014), noise-vocoded sentences (Neger et al., 2014) nor speech in noise (Sommers & Danielson, 1999). However, conflicting results have been reported for older listeners. Sommers and Danielson (1999) reported that processing speech did not influence older listeners' speech perception abilities of high and low predictability stimuli in noise. In contrast, Neger et al. (2014) found that older listeners' processing speed predicted their perception of noise-vocoded sentences. Other studies have also found effects of processing speed on listeners' speech recognition abilities. For example, Woods et al. (2013) found increased processing speed positively correlated with sentence in noise identification for a wide range of normal hearing listeners (aged 19-61), and aided listeners with hearing impairment (aged 31-73). Furthermore, Desjardins and Doherty (2013) found that increased processing speed was predictive of improved speech recognition in noise for both their groups of older and younger adults. Another study (Humes, Kidd, & Lentz, 2013) including processing speed in a composite score of cognition also found that cognition accounted for 11.4% of the variance in older listeners' (aged 60 – 86 years, M = 69.2 years) aided speech understanding.

Non-verbal intelligence. Non-verbal intelligence has also been included in previous speech recognition studies in order to account for listeners' general intelligence without the bias of language. Earlier studies that attempted to link auditory discrimination and individual listener differences used general intelligence as a predictor (e.g., Spearman, 1904; Watson, 1991). Since then, however, more recent studies chose to focus on narrower areas of cognition with processes thought to pertain more directly to speech recognition, e.g., working memory and processing speed. In narrowing the focus, it is still necessary to rule out the possibility that improved speech understanding is simply a by-product of higher IQ. Two recent studies showed no link between non-verbal intelligence and speech perception in noise. Tamati et al. (2013) found that differences in younger listeners' speech recognition in noise accuracy was not predicted by non-verbal intelligence, as measured by the WASI Performance IQ subtests of Block Design and Matrix Reasoning (Wechsler, 1999). This finding was in line with their prediction that the ability to recognise highly variable stimuli in adverse conditions is specific to speech processing and not to general intelligence. Conway et al. (2010) also failed to link the ability to use sentence context in spectrally degraded speech and non-verbal intelligence. In addition, Conway et al. (2010) found that statistical learning was significantly correlated with the ability to use sentence context in the perception of spectrally degraded speech. Statistical learning is the ability to implicitly determine structure in the input by detecting the probabilities of its items co-occurring (Misyak & Christiansen, 2012). Given that language is by nature probabilistic (Auer & Luce, 2005), a number of studies have begun to link statistical learning with language. Statistical learning has been shown to be important for language acquisition (Saffran, 2003) and in adult sentence comprehension (Misyak & Christiansen, 2012) , yet the underlying mechanisms have only received limited investigation.

In spite of listeners' general intelligence not predicting their speech recognition abilities (Conway et al., 2010; Tamati et al., 2013), the findings that statistical learning is a predictor of speech recognition (Conway et al., 2010) suggests that listeners utilise other processes to aid speech understanding in adverse conditions.

Vocabulary knowledge. Vocabulary knowledge refers to the number of words present in a language user's lexicon. Vocabulary knowledge is also often considered a proxy measure of language experience; since older adults and native speakers commonly exhibit higher vocabulary than younger listeners (e.g., McAuliffe et al., 2013; Neger et al., 2014; Sheldon, Pichora-Fuller, & Schneider, 2008) and non-native speakers (e.g., Tamati & Pisoni, 2014) respectively. Recently, the potential link between a listener's vocabulary knowledge, or language experience, and their speech perception abilities has been highlighted by a number of authors (e.g., Benard et al., 2014; McAuliffe et al., 2013; Tamati et al., 2013). It appears that vocabulary knowledge may be another factor that can further explain individual differences in adverse listening conditions.

Interestingly, some studies have examined vocabulary and working memory as predictors of perception within the same sample—and reported that vocabulary knowledge, not working memory, was linked to speech recognition in adverse conditions. McAuliffe et al. (2013) found that a listener's receptive vocabulary, as measured by the Peabody Picture Vocabulary Test (PPVT-IV) (Dunn & Dunn, 2007), predicted the perception of moderately degraded speech (i.e., hypokinetic dysarthria, a naturally degraded speech signal associated with Parkinson's disease). For younger listeners (M= 20 years), vocabulary was the only significant predictor of intelligibility; in older listeners (M = 65 years), both hearing sensitivity and vocabulary significantly affected intelligibility. Older listeners' vocabulary knowledge benefited speech recognition until hearing impairment began to impact on speech perception. In a similar vein, Neger et al.'s (2014) younger listeners' perceptual learning of noise-vocoded speech was not predicted by working memory or processing speed; however, younger listeners' perceptual learning was modified by vocabulary knowledge. That is, younger listeners with larger vocabularies also exhibited increased understanding of noise-vocoded speech over time.

Other studies have also found links between vocabulary knowledge and speech recognition for a variety of listener populations and adverse conditions. Alamsaputra, Kohnert, Munson, and Reichle (2006) found that the PPVT-III (Dunn & Dunn, 1997) was the only measure of linguistic experience to predict non-native English speakers' ability to identify synthesised speech. The researchers failed to find an effect for Test of English as a Foreign Language (TOEFL) scores nor length of US residency (1.5-11years) as predictors of sentence identification performance. Munson (2001) also reported children with larger expressive and receptive vocabularies were better at identifying degraded CVC words than children with smaller vocabularies. Janse and Adank (2012) showed that older listeners with a higher receptive vocabulary adapted to a novel accent quicker than older listeners with lower receptive vocabulary scores. Tamati et al's (2013) good listeners reported higher word familiarity than their poor listeners on a measure of self-reported vocabulary knowledge (WordFam (Pisoni, 2007)). Furthermore, Benard et al. (2014) identified that vocabulary knowledge as assessed by the PPVT-III (Dunn & Dunn, 1997) correlated with listeners' abilities to identify interrupted speech.

However, not all studies have found links between vocabulary knowledge and speech recognition in adverse conditions. Tamati and Pisoni (2014) failed to find a link between non-native speakers self-reported vocabulary knowledge and sentence recognition. The authors suggested that the listeners' level of vocabulary knowledge may have been too low and too invariable to demonstrate an effect. In comparison to native speakers, non-native speakers provided lower familiarity ratings for words rated as highly familiar by native speakers. Furthermore, they suggested that listening conditions may have been too adverse for participants to benefit from their limited vocabulary knowledge. Benichov, Cox, Tun, and Wingfield (2012) also did not find expressive verbal ability a predictor of word recognition in noise for older and younger listeners for a range of sentence contexts.

In sum, the literature on vocabulary influences on speech perception is somewhat inconsistent. However, the studies above which found effects of vocabulary mostly used receptive vocabulary tasks (Alamsaputra et al., 2006; Benard et al., 2014; Janse & Adank, 2012; McAuliffe et al., 2013; Neger et al., 2014; Tamati et al., 2013). On the other hand expressive vocabulary, was a predictor for children's word recognition (Munson, 2001), yet Benichov et al. (2012) found no effect with a productive vocabulary task, the WAIS-IV vocabulary subtest (Wechsler, 2008) for adult listeners.

Summary of individual difference findings. Recent studies have attempted to pinpoint which individual differences might be related to performance—yet findings to date have been equivocal. However, many of the studies investigating cognition and speech understanding include older listeners and listeners with hearing loss (e.g., Krull et al., 2013; Woods et al., 2013). Furthermore, studies including listeners with hearing loss vary as to whether listeners are aided (e.g., Humes et al., 2013) or unaided (e.g., Neger et al., 2014). Nevertheless, younger normal hearing listeners also demonstrate variation in their ability to accurately recognise speech in adverse conditions (e.g., Gilbert et al., 2013). Yet, few of the studies reviewed considered the cognitive abilities of their younger normal hearing listeners (e.g., McAuliffe et al., 2013; Tamati et al., 2013; Woods et al., 2013). Thus, it is difficult to remove the confounding effects of age and hearing sensitivity in order to establish underlying cognitive effects on speech recognition in difficult listening environments. Studies investigating vocabulary, on the other hand, have often focused on listeners with limited language experience, non-native speakers (e.g., Alamsaputra et al., 2006) and children (e.g., Munson, 2001), yet healthy younger listeners also exhibit differences in vocabulary knowledge in comparison to each other (Tamati et al., 2013) and to older listeners (McAuliffe et al., 2013; Neger et al., 2014). These differences suggest further investigation of vocabulary knowledge's influence on speech recognition in younger listeners with normal hearing is warranted.

In addition to these listener differences, the studies reviewed have also varied with respect to the ecological validity of their stimuli. Some studies used interrupted speech (Benard et al., 2014) or frequency compressed speech (Ellis & Munro, 2013)—situations not naturally encountered by listeners. This also may explain why the individual factors of processing speed and working memory were not predictors of speech recognition in those studies. Another way to maintain ecological validity, as suggested by Gilbert et al. (2013), is to have highly variable stimuli. Every day listeners encounter a myriad of different talkers. Thus, including stimuli with multiple male and female talkers is important.

Another way in which the studies reviewed vary is whether they use sentences (e.g., Benard et al., 2014; Tamati et al., 2013), or isolated words (e.g., Krull et al., 2013; Munson, 2001) as stimuli. Within studies that investigate sentences, some utilise high context stimuli (e.g., Alamsaputra et al., 2006; Benard et al., 2014), others use low context stimuli (e.g., Ellis & Munro, 2013), and some utilise both (e.g., Desjardins & Doherty, 2013; Humes et al., 2013). Similar effects of vocabulary and cognition on speech recognition have been found for both high and low context stimuli. For example, Desjardins and Doherty (2013) found positive effects of both increased working memory and processing speed on word recognition accuracy for both their high and low context sentences. Increased vocabulary knowledge has also been shown to be a predictor of speech recognition for both low (e.g., McAuliffe et al., 2013) and high context stimuli (e.g., Neger et al., 2014). The absence of an effect of non-verbal intelligence on speech recognition appears to be consistent for both high and low context sentences (e.g., Conway et al., 2010). Isolated word stimuli, on the other hand, provide less consistent results for the effects of vocabulary and cognition on word recognition accuracy. Isolated word stimuli in the current review have not shown an effect for working memory (Krull et al., 2013), nor processing speed (Sommers & Danielson, 1999), but did so for vocabulary (Munson, 2001). However, the number of studies utilising isolated words is limited. The current study employs non-sensical phrases, in keeping with previous studies from this laboratory (McAuliffe et al., 2013; McAuliffe, Kerr, Gibson, Anderson, & Lashell, 2014), to reduce top-down semantic influences on speech recognition.

The current study aims to address the highlighted issues by investigating a number of individual difference measures (working memory, processing speed, non-verbal intelligence, and vocabulary knowledge) in relation to younger normal hearing listeners' speech understanding accuracy for multiple talker stimuli in noise. Nevertheless, listener differences are not the only differences which have the ability to predict speech recognition—lexical differences in the language may also provide information about the underlying mechanisms of speech recognition in adverse conditions. The subsequent section changes focus and discusses lexical differences in the language as another factor that may influence speech recognition.

Lexical Cues and Speech Recognition

A number of researchers have taken a language-centred approach to speech recognition and investigated underlying lexical factors in the language as predictors of speech understanding. Two important lexical cues which have been shown to play a role in speech recognition are phonological similarity and lexical frequency.

Phonological similarity is often measured using neighbourhood density. A word's neighbourhood contains all the similar sounding words that are created by either adding, deleting, or substituting a single phoneme to the target word (e.g., Luce & Pisoni, 1998). Neighbourhood density has been suggested to relate to lexical competition because word recognition involves discriminating among similar sounding lexical items (neighbours) in memory (Luce & Pisoni, 1998). Words with a large number of neighbours (high neighbourhood density) have more competitors, and are therefore more difficult to recognise than words with fewer neighbours (McArdle & Wilson, 2008; Vitevitch & Luce, 1998, 1999).

Lexical frequency is the number of times a word is used in language. Numerous studies have reported that high frequency lexical items are perceived and produced quicker than low frequency lexical items (e.g., Balota & Chumbley, 1984; Jescheniak & Levelt, 1994; Monsell, Doyle, & Haggard, 1989), as well as more accurately (e.g., Goldinger, Luce, & Pisoni, 1989; Howes, 1957). Furthermore, lexical frequency also works together with neighbourhood density to reduce competition, by biasing the system towards the higher frequency words (McArdle & Wilson, 2008).

Studies investigating normal hearing and hearing impaired populations have provided support for the hypotheses that neighbourhood density and word frequency play a role in speech recognition. Dirks, Takayanagi, Moshfegh, Noffsinger, and Fausti (2001) and Dirks, Takayanagi, and Moshfegh, (2001) investigated the effects of lexical neighbourhoods and word frequency on word recognition for normal hearing listeners and listeners with sensorineural hearing loss. Dirks et al. (Dirks, Takayanagi, & Moshfegh, 2001; 2001) concluded that both normal hearing and hearing impaired listeners recognised high word frequency, low density neighbourhood and low average frequency neighbourhood items (easy words) more accurately than items with low frequency, high density and high average frequency neighbourhoods (hard words) in both quiet and noisy conditions. Takayanagi, Dirks, and Moshfegh (2002) found easy words were recognised more easily than hard words for all of their listener groups (native normal hearing, native hearing impaired, non-native normal hearing, and non-native hearing impaired listeners).

More recently, Taler, Aaron, Steinmetz and Pisoni (2010) investigated the effect of lexical frequency and neighbourhood density on normal hearing older and younger adults' word recognition accuracy. Overall, high frequency words were recognised more accurately than low frequency words. Low neighbourhood density items were recognised more accurately than high neighbourhood density items. These frequency and neighbourhood density effects were stronger at a lower signal to noise ratios, and neighbourhood density effects were stronger in phrases with low semantic predictability. Furthermore, the neighbourhood density effects were stronger for older adults than younger adults. These age differences were particularly strong when listening conditions were most difficult, i.e. low semantic probability, low frequency and low signal-to-noise ratio (SNR) conditions. These findings are consistent with the previous reports of frequency and neighbourhood density effects; however, they also suggest that listening conditions are important in determining the extent of linguistic cue use.

Other studies have opted to use non-word stimuli in order to control lexical cues. For example, Janse and Newman (2013) investigated the effect of neighbourhood densities on non-word identification. Importantly, the effect of neighbourhood density for non-words is facilitatory rather than inhibitory; as the neighbourhood density increases, non-words are easier to identify. This facilitatory effect reduced generalisation to real-world listening conditions.

However, not all studies have found the expected effects of neighbourhood density and lexical frequency. McArdle and Wilson (2008), for example, found no effect of neighbourhood density, or lexical frequency as predictors for monosyllabic word speech reception thresholds for normal hearing listeners. The authors suggested that lack of variance in the sample's word frequencies may have contributed to their findings. In contrast, Freedman and Barlow (2013) found that words with higher neighbourhood densities were recognised more accurately than words with lower neighbourhood densities. To directly measure the effects of neighbourhood density on word recognition, Freedman and Barlow (2013) controlled their stimuli for lexical frequency. However, since there was no difference in frequency for the high and low neighbourhood density word groups, lexical frequency effects were not investigated. Nevertheless, controlling for lexical frequency is not uncommon. For example, Vitevitch and Luce (1998, 1999) also controlled for lexical frequency in order to isolate the effects of neighbourhood density on word recognition.

Neighbourhood density is not the only measure of phonological similarity available. Phonological Levenshtein distance (PLD) provides an alternative measure of lexical similarity. Levenshtein distance (LD) is used in computer science in order to quantify the similarity of strings and has practical application in spell checkers. LD is defined as the number of insertions, deletions, and substitutions needed to convert one string of elements (e.g., letters or phonemes) to another. Orthographic LD measures have been shown to account for more variance in lexical access tasks than standard orthographic neighbourhood measures (Yarkoni, Balota, & Yap, 2008). Furthermore, Yap and Balota (2009) demonstrated phonological LD's utility as a measure of word similarity for long words. They found that PLD accounted for differences in multisyllabic word recognition whereas neighbourhood density could not. The authors highlighted that many long words have no phonological neighbours. However, these long words are not completely dissimilar from other words; PLD is able to capture this similarity. Thus, PLD is a more fine-grained analysis of phonological similarity than neighbourhood density.

Nevertheless, to the researcher's knowledge, only one study (Suárez, Tan, Yap, & Goh, 2011) has investigated PLD in relation to spoken word recognition. Suárez et al. (2011)

investigated speech recognition of lexical hermits (words with neither traditional phonological nor orthographic neighbours), in order to demonstrate the utility of PLD as a measure of lexical similarity. The authors found that PLD provided evidence of lexical competition. Low frequency words that were more similar to other words (lower PLD) were recognised less accurately than words that were more phonologically distant (higher PLD). As Suárez et al. (2011) chose to use lexical hermits, unfortunately their results can only be directly compared to previous studies' findings for words with low neighbourhood densities.

To summarise, although studies investigating lexical characteristics as predictors of speech recognition demonstrate more consistent results than studies investigating individual characteristics, not all studies have found lexical characteristics to be predictors of speech recognition. It appears possible that these inconsistent findings may have resulted from methodological differences.

Many previous studies have used neighbourhood density as a measure of phonological similarity; however, there are also a large number of words in the lexicon which have no phonological neighbours (Vitevitch, 2008). Nevertheless, these so called lexical hermits have lexical properties which influence speech recognition (Suárez et al., 2011). The present study extends upon Suárez et al.'s (2011) work by utilising PLD as its measure of phonological similarity for real word stimuli.

Moreover, in order to investigate the effects of lexical cues it is necessary to reduce contextual cues. By reducing contextual cue use, listeners are able to focus on the piecing together of glimpsed material without the added benefit of semantic knowledge (Krull et al., 2013). Thus, the current study aims to address these issues by employing the use of non-sensical phrases.

Lastly, many previous studies of linguistic properties have involved non-words (Janse & Newman, 2013; Storkel, 2013; Storkel, Armbruster, & Hogan, 2006). However, Storkel (2013) compared non-word and real word phonotactic and neighbourhood density properties for all possible legal CVC combinations and found significant differences between phonotactic probabilities and neighbourhood density measures for non-words and real words, reducing the ecological validity of non-words, which in turn reduces generalisation to real word stimuli. Furthermore, neighbourhood density for non-words also provides facilitatory effects and not competition, which further reduces the generalisation of results to everyday situations. Thus, it is important to make use of real word stimuli where possible to maintain ecological validity. Therefore, although the current study utilises non-sensical phrases, the words within the phrases are all real English words.

Although the effects of listener characteristics and lexical cues on speech recognition have been investigated in numerous studies in isolation, the underlying mechanism of the role of individual differences in speech recognition is still unknown. The next section explores studies which have investigated these listener characteristics and lexical cues together in order to further understand speech recognition in adverse conditions.

The Interaction between Individual Differences and Language-Based Factors

Few studies have investigated the link between listener characteristics (e.g., hearing impairment, working memory and processing speed) and the use of lexical cues (e.g., phonological similarity and word frequency) to aid speech recognition in adverse conditions (Bradlow & Pisoni, 1999; Frisch, Large, Zawaydeh, & Pisoni, 2001; Janse & Newman, 2013; McAuliffe et al., 2013; Sommers & Danielson, 1999; Taler et al., 2010). These studies have for the most part only investigated one or two listener characteristics and lexical cues; however, their findings suggest that these interactions play an important role in speech understanding in difficult listening conditions.

Some studies have reported links between listener characteristics and lexical cue use. For example, Sommers and Danielson (1999) noted that inhibition, based on a combined measure of the ability to inhibit semantic, phonetic and voice information, was significantly related to participants' identification scores for high density words for sentences with both high and low predictability, i.e., participants with poorer inhibition had greater difficulty identifying high density words. Similarly, Taler et al. (2010) found that poor working memory, as measured by a Stroop naming task and an elevated better ear 2kHz threshold both correlated with poorer identification of high neighbourhood density (more difficult) words in the poor SNR condition (-3 dB SNR). The researchers also found that better short term memory, as measured by the forward digit span, was significantly correlated with more accurate identification of high frequency words at the higher SNR (+10 dB SNR). Furthermore, although Janse and Newman (2013) found that younger adults with poor attention switching control generally performed worse at identifying non-words, they also benefited more from the facilitatory effects of neighbourhood density. Older adults with poorer hearing also benefited more from increasing neighbourhood density effects.

However, listener characteristics do not always have an effect on lexical cue use. Unlike Taler et al. (2010), Janse and Newman (2013) did not find an effect for short term nor working memory interacting with neighbourhood density and speech recognition of non-words. Processing speed has also failed to provide an insight into neighbourhood density use for normal hearing listeners' understanding of speech in noise (Sommers & Danielson, 1999) nor for older adults identification of non-words in noise (Janse & Newman, 2013).

Given that lexical characteristics are inherent to the language, a logical step would be to investigate the influence of listeners' lexical knowledge on the use of lexical cues of phonological similarity and lexical frequency in speech recognition. However, only a few studies have investigated this link. McAuliffe et al. (2013) ran post-hoc analysis to exclude possible lexical frequency effects on vocabulary knowledge's ability to predict dysarthric speech understanding. They found no main effect nor interactions with phrasal frequency nor the lowest frequency item in the phrase for any of their individual measures. They concluded that—in their sample—the vocabulary knowledge effect was not moderated by lexical frequency.

Whereas McAuliffe et al. (2013) investigated the link between vocabulary and lexical frequency, Frisch et al. (2001) examined the link between vocabulary knowledge and neighbourhood density. Their study found no link between neighbourhood density, vocabulary knowledge and the well-formedness judgements of non-words. Listeners with higher word familiarity ratings for medium and low familiarity words (based on a shortened version of the FAM test (Lewellen, Goldinger, Pisoni, & Greene, 1991)) were more likely to judge non-words as well formed. That is, they were more likely to accept a non-word with a lower probability of occurrence for its phonemes. Given that in general the word familiarity items in the FAM test were highly correlated with their occurrence in the language, Frisch et al. (2001) suggested that a lexicon with more lower frequency lexical items (a larger lexicon) would contain more similar items to the low probability non-words in their corpus. However, in their simulation they found that low probability non-words had no lexical neighbours in neither their large lexicon (all words in Webster's Pocket Dictionary were potential neighbours) nor their small lexicon (words which were rated as 6 or higher in the FAM, based on Nusbaum, Pisoni, & Davis' (1984) normative data). Thus, Frisch et al. (2001) concluded that the link between word familiarity and the well-formedness judgements of non-words are not accounted for by differences in lexical neighbourhoods.

To the researcher's knowledge, only one study (Bradlow & Pisoni, 1999) has investigated vocabulary knowledge, neighbourhood density and word frequency together. Bradlow and Pisoni (1999) reported that both native and non-native speakers of English exhibited higher recognition accuracy for easy words (low neighbourhood density, high frequency) compared to hard words (high neighbourhood density, low frequency); however, the non-native speakers showed a much larger difference in word recognition between the easy and hard words. Native and non-native speakers also exhibited differences in word familiarity. Non-native speakers rated easy words as more familiar and hard words as less familiar, whereas native listeners rated easy and hard words with the same high familiarity. Even when only highly familiar words were considered, easy words were still recognised more readily than hard words by non-native speakers.

In order to explain the differences between native and non-native speakers' familiarity and their recognition of easy and hard words, the authors ran a series of correlational analyses between these scores and demographic factors. None of the demographic factors (age of English study onset, number of years studying English, and number of years spent in an English environment) correlated with the easy words. For hard word recognition, number of years in an English environment was positively correlated with word accuracy. For hard word familiarity, age of English study onset was negatively correlated with familiarity rating. Based on these results, the researchers suggest that spoken word recognition requires exposure to spoken input, whereas written vocabulary, linked to word familiarity, benefits most from beginning formal language study earlier. Thus, hard word familiarity and hard word recognition appear to indicate different aspects of language proficiency. Yet, these listener factors were not directly used as predictors of lexical cue use which limits generalisations, and suggests further investigation is necessary.

Similarly, Neger et al.'s (2014) findings also suggest that vocabulary knowledge links with different areas of language ability. The authors reported that both higher vocabulary listeners and listeners who did well in a statistical learning task, also did well in the perceptual learning task. This finding suggests that high vocabulary listeners are tapping into certain other underlying cues in the language and using them to aid perception, yet what these cues are is currently unclear.

In summary, a limited number of studies have investigated individual differences and lexical characteristics together as predictors of word recognition. These studies focused on a small number of individual differences only, with varying success at identifying predictors of lexical cue use in speech recognition. Only two studies found a link between individual differences and lexical cue use. Taler et al. (2010) found links between working memory, hearing impairment and neighbourhood density, as well as between short term memory and high lexical frequency. However, in the other study to find a link between individual differences and lexical cue use (Janse & Newman, 2013) generalisation is limited to non-words. Furthermore, studies which have included vocabulary as a predictor have failed to directly pit vocabulary knowledge against lexical cue use, with the exception of McAuliffe et al. (2013) who investigated phrasal frequency only. Thus, the current study aims to directly investigate the interactions between vocabulary knowledge and cognitive measures and the lexical cues of phonological similarity and lexical frequency as predictors of word recognition in adverse conditions.

Aims and Hypotheses

The current study aims to investigate the link between individual listener differences, in particularly vocabulary, and the use of the lexical cues of phonological similarity and lexical frequency during speech recognition in adverse listening conditions for younger listeners with normal hearing. To isolate effects of vocabulary from other influences such as cognitive ability and hearing loss, measures of working memory, processing speed, non-verbal intelligence, puretone audiometry, and general cognition were also included. In including this large range of possible predictors, the purpose is to gain an insight into the underlying mechanisms of speech recognition in adverse listening conditions.

From these general aims more specific research hypotheses arise:

1. With respect to adverse conditions in general, it is hypothesised that as listening conditions become more favourable (SNR increases), listeners' word recognition accuracy will increase, (e.g., Miller, 1947; Miller et al., 1951; Plomp, 1978).
2. It is hypothesised that individual vocabulary and cognitive measures will show variance in their ability to predict word recognition in adverse conditions:

    a. Listeners with larger receptive vocabularies will exhibit significantly higher word recognition accuracy than listeners with smaller vocabularies, (e.g., Alamsaputra et al., 2006; Benard et al., 2014; McAuliffe et al., 2013).
    b. Working memory will be significantly positively related to speech recognition. In particular, listeners with increased working memory will demonstrate higher word recognition accuracy (e.g., as per Akeroyd, 2008; Tamati et al., 2013).
    c. Given this sample includes younger normal hearing listeners only, there will be no relationship between processing speed and speech recognition (e.g., Benard et al., 2014; Neger et al., 2014; Sommers & Danielson, 1999)
    d. Likewise, non-verbal intelligence is not expected to influence speech recognition (e.g., Tamati et al., 2013).
3. With respect to lexical characteristics, it is expected that lexical frequency and phonological similarity will have different effects.
    a. Higher frequency items will be more readily recognised than lower frequency items (e.g., as per Goldinger et al., 1989; Howes, 1957; Taler et al., 2010).
    b. Whereas, words with increased phonological similarity to other words will be less accurately identified due to their increased confusability (e.g., as per Luce & Pisoni, 1998; Suárez et al., 2011; Vitevitch & Luce, 1998, 1999).

4. The extent to which listener characteristics and language-based factors interact is largely unexplored. However, based upon the preceding review the following hypotheses are put forward.
    a. Increased working memory will result in better identification accuracy for higher frequency items (e.g., as per Taler et al., 2010).
    b. Poorer working memory is expected to result in poorer identification of items with increased phonological similarity to other words (e.g., as per Taler et al., 2010).
    c. Non-verbal intelligence and processing speed are not expected to interact with lexical characteristics, given their lack of predictive power in isolation (e.g., Benard et al., 2014; Neger et al., 2014; Sommers & Danielson, 1999; Tamati et al., 2013) and lack of relation with lexical cues (e.g., Janse & Newman, 2013; Sommers & Danielson, 1999).
    d. A relationship is expected between vocabulary and lexical cues, but the direction of that relationship is not predicted.

Method

Participants

Fifty listeners (40 females and 10 males) were recruited for this study as part of a larger project (McAuliffe & Sinex, funded 2014-2017). All listeners were native speakers of New Zealand English and aged 18 to 33 years (M age = 21.46 years, SD = 2.96). Furthermore, listeners reported no language, learning, or cognitive disabilities, and showed no evidence of cognitive impairment on the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005). All listeners exhibited behavioural air conduction puretone thresholds within normal limits bilaterally. Specifically, all thresholds were less than or equal to 20 dB HL (Schlauch & Nelson, 2009). Puretone thresholds were measured for octave frequencies 0.25 – 8 kHz using a GSI 61 two-channel audiometer with Telephonics TDH-SDP supra-aural headphones in a soundproof booth. Listeners were recruited via university, social media, researchers' friends and family, and prior participant databases. Approval for the study was obtained from the University of Canterbury Human Ethics Committee. Each participant provided both verbal and written consent to participate.

Method Overview

This experiment comprised two primary components: (1) individual measures of vocabulary and cognition and (2) a speech perception task. Each of these components is outlined in the sections below.

Individual measures of vocabulary and cognition. Each participant completed a series of assessments that aimed to evaluate the role of individual differences in speech perception. These tasks included assessments of receptive vocabulary, working memory, non-verbal intelligence, and processing speed. The individual measures are outlined in subsequent sections.

Vocabulary knowledge. Vocabulary was assessed using the Peabody Picture

Vocabulary Task (PPVT-IV) (Dunn & Dunn, 2007). The PPVT is a standardised measure of receptive vocabulary commonly used to assess verbal ability (e.g., Alamsaputra et al., 2006; Benard et al., 2014; McAuliffe et al., 2013). The test consisted of nineteen sets of twelve items and was administered according to standard testing procedures. The tester read aloud each item and participants selected which one of four pictures presented best represented the item. Raw scores were converted to standard scores prior to data analysis.

Working memory. Working memory was assessed using the digit span subtest of the Wechsler Adult Intelligence Scale (WAIS-IV) (Wechsler, 2008), and a reading span (RSPAN) task adapted from Tompkins, Bloise, Timko, and Baumgaertner (1994).

The digit span section comprised of three tasks that measured forward, backward and sequencing digit spans. The digit span tasks were administered as per the WAIS-IV manual (Wechsler, 2008). Digit span forward required participants to repeat the items in the same order as presented; for digit span backward, participants recalled the numbers in reverse order; and for digit span sequencing, participants stated the numbers in order from lowest to highest. Each span is scored out of 16, for a possible total raw score of 48.

The RSPAN task was included as a measure of verbal working memory. This task, with its increased language emphasis, was selected to complement the digit span task as it is often shown to correlate with speech reception (Akeroyd, 2008). The task was programmed in DirectRT (Jarvis, 2010). Participants were presented with on screen written instructions and practice stimuli prior to beginning the task. The task consisted of 12 sets of sentences (three sets of two sentences, three sets of three sentences, three sets of four sentences, and three sets of five sentences) for a total of 42 sentences. Participants were told that the number of sentences would increase without warning. Sentences were presented one at a time on screen. Participants were asked to read aloud each sentence, say whether the sentence was “true” or “false”, and remember the final word of the sentence. They did this for all sentences in the set, until an “X” appeared on the screen, and then they were to recall all the sentence final words they could remember for that set. Participants were asked to repeat the words in the same order that they occurred. However, if this was not possible, they were asked not to recall the last word first. The tester controlled the presentation of the sentences. Once the participant had indicated the veracity of the sentence, the tester presented the next sentence. This method was used in order to control the time participants took to complete the task, and reduce the chance for the participant to employ learning strategies (Friedman & Miyake, 2004). The RSPAN score was the total number of words recalled (maximum = 42).

Non-verbal intelligence. In line with previous research (Tamati et al., 2013), two subtests of the WAIS-IV (Wechsler, 2008), block design and matrix reasoning, were included to investigate participants’ non-verbal intelligence**.** Both tasks were administered as per WAIS-IV protocols (Wechsler, 2008).

The block design task involved recreating a picture with red and white blocks within a time limit. Participants received a raw score out of 66. For the matrix reasoning subtest, the participant selected the response which completed the matrix or series from five options. There were two sample items which demonstrated the two trial types in this subtest: 2 X 2 matrix and series items. Each correct item received a score of 1, for a total possible raw score of 26.

Processing speed. Processing speed was assessed to account for individual differences in ability to efficiently perform simple operations (Kaufman et al., 2010). Previous studies of speech perception (Benard et al., 2014; Desjardins & Doherty, 2013; Neger et al., 2014) have employed processing speed measures from the WAIS (Wechsler, 1997a, 2008). Two pen and paper subtests of the WAIS-IV (Wechsler, 2008), symbol search and coding, were used to measure processing speed. Both tasks were administered as per WAIS-IV protocols (Wechsler, 2008).

The symbol search subtest required participants to search for symbols from a target group inside a search group within a time limit. A total raw score of 60 was possible. The coding subtest involved participants copying symbols paired with numbers within a time limit. A total raw score of 135 was possible.

The previous sections outlined the vocabulary and cognition measures utilised in the study; the speech perception task will be detailed in the following sections.

Speech perception stimuli. Experimental stimuli for the speech perception task were selected in accordance with standard lab procedures. However, for the current study, further investigation of word frequency and word similarity characteristics of the stimuli was conducted as part of the initial phrase selection. Subsequently, all phrases were recorded and a detailed protocol of stimuli selection, counterbalancing and blocking was devised. These procedures are detailed below in three sections: (1) linguistic composition of experimental stimuli, (2) stimuli recording procedures, and (3) stimuli selection and counterbalancing procedures.

Linguistic composition of experimental stimuli. Speech stimuli consisted of 128 semantically anomalous phrases (see Appendix). The 128 phrases had alternating stress contrasts (64 strong-weak and 64 weak-strong), as per previous studies from this laboratory (McAuliffe et al., 2013, 2014). Each phrase was six syllables long and contained between three and five words. In order to lend themselves to the study of lexical cue use for our listener population, these 128 phrases were adapted from a pool of 160 phrases to ensure that no duplication of content words occurred and that all phrases were consistent with New Zealand English. In addition, measures of lexical frequency and phonological similarity were obtained from the English Lexicon Project (ELP) (Balota et al., 2007) for each of the 358 content words in the phrases. The ELP is a collaboration between six North American universities which provides access to a dataset of lexical properties and behavioural data (visual lexical decision and naming tasks) for 40,481 words and 40,481 non-words. The ELP can be accessed at elexicon.wustl.edu. A number of different measures exist to examine lexical frequency and phonological word similarity, each with their associated strengths and weaknesses. The current study utilised two measures provided by the ELP: the log Hyperspace Analogue to Language (HAL) frequency norms (Lund & Burgess, 1996) as a measure of word frequency; and phonological Levenshtein distance (PLD) as a measure of phonological similarity. The rationale for choosing these two measures is outlined below.

Lexical frequency. As stated, lexical frequency was calculated based on the log Hyperspace Analogue to Language (HAL) frequency norms (Lund & Burgess, 1996). These norms are based on the HAL corpus, which consists of 131 million words gathered from 3000 Usenet newsgroups in Februrary 1995 (Lund & Burgess, 1996)[1]. HAL frequency norms have been shown to account for more variance in accurate identification and lexical decision times of both mono- and bi-syllabic words than other corpora, such as the Kučera and Francis (1967) frequency norms and CELEX (Baayen, Piepenbrock, & van Rijn, 1993), commonly cited in the literature. Furthermore, Brysbaert and New (2009) state that a corpus smaller than 16 million (e.g., Kučera and Francis (1967) frequency norms) is not reliable enough to estimate frequencies for words with frequencies lower than 10 per million. Thus, the HAL frequency norms were selected as a suitable measure of lexical frequency. Mean lexical frequency for the 358 content words was 9.91 (SD = 1.42, range = 5.49-13.58).

Phonological similarity. Phonological similarity was calculated using phonological Levenshtein distance (PLD). Traditionally phonological similarity is measured by neighbourhood density. However, PLD provides another measure of phonological similarity. PLD was chosen because of a number of reasons: it is a continuous measure of phonological

---

[1] According to the English Lexicon Project, the HAL frequency estimates are currently based on approximately 400 million words, rather than the original 131 million reported by Lund and Burgess (1996).

similarity; the largest part of the lexicon has no phonological neighbours (Vitevitch, 2008); and the only study of PLD and spoken word recognition to date (Suárez et al., 2011) has limited generalisability in that they only investigated lexical hermits.

PLD measures in the ELP are based on phonological transcriptions from the Unisyn Lexicon (Unisyn lexicon, n.d.). An individual word's PLD is calculated firstly by measuring the LD between the target word's phonological transcription and every other word's phonological transcription in the ELP. Next, the mean LD of the target word's 20 nearest phonological neighbours is calculated. Thus, a word's PLD is the average of the 20 words with the shortest LD in relation to the target word phonology (Yap & Balota, 2009). Words with smaller PLDs have more phonological similarities with other words. Mean PLD for the 358 content words was 1.78 (SD = 0.54, range =1-3.85).

Stimuli recording procedures. All 128 phrases were spoken by eight healthy New Zealand English talkers (four male and four female) and recorded in line with existing laboratory procedures. Talkers were seated in a sound proof booth and instructed to read each phrase from a computer screen aloud twice in their everyday speaking voice. Each phrase was presented on the screen via PowerPoint presentation for ten seconds. Before the talkers read aloud the phrases, they heard four of the phrases modelled by another talker in order to cue a suitable speech rate. Talkers could take a short break after each set of 32 phrases and the model phrases were repeated again.

Monaural digital audio recordings (44.1 kHz sampling rate, 16 bits quantisation) were captured directly to a compact flash memory card via an Earthworks M30 desk microphone situated 30cm to the side of the talker coupled to a TASCAM HD-P2 portable stereo recorder. Recordings were carefully monitored with talkers asked to repeat any phrases which contained errors

Stimuli selection and counterbalancing procedures. Firstly, the single recording for each talker was manually segmented into individual files for each phrase using MATLAB (Mathworks, n.d.).The best example of each phrase was chosen for use in the speech perception task. This involved selecting the phrase which sounded most natural. If there was no perceptible difference then the first instance of each phrase was chosen. Next, the selected audio recordings were mixed with noise matched to the long term average speech spectra of phrases produced by the same talker to obtain different levels of degradation (signal-to-noise ratios, SNRs), as per Sinex (2013). The level of speech was always 65 dB A and the noise level was varied, as per Gilbert et al. (2013).

Pilot selection of degradation levels. A short pilot study was conducted in order to confirm SNRs which would elicit a range of phrase accuracy scores. Pilot participants were five adults who reported no hearing difficulties. Each pilot listener heard all 128 phrases (randomised and counterbalanced as outlined in the following section) divided into 32 phrases at four SNRs selected from 14 SNRs ranging from -7 to +6 dB SNR. From these 14 SNRs, four (-5, -2, +1 and +4 dB SNR) were chosen for the final speech perception experiment, as they elicited on average between 28% and 83% words correct per phrase scores. These SNRs were chosen to avoid floor and ceiling effects, and to allow for systematic investigation of SNR level on word accuracy.

Counterbalancing and randomisation of speech stimuli. For each listener, a speech stimuli list was generated which contained each of the 128 phrases, divided into four counterbalanced blocks of 32 phrases per each of the four dB SNR levels. The lists were balanced so that at each dB SNR level listeners heard four phrases spoken by each of the eight talkers (four female and four male). Within each SNR, each talker's phrases were balanced to include one of higher average word frequency and one of lower average frequency for each stress pattern (strong-weak and weak-strong). Furthermore, average frequency of the phrase and phrase length were balanced within each talker by stress pattern group, as well as across stress pattern and across speakers. Lists were generated and chosen by random selection and rejected if they did not meet balancing criteria.[2]

Procedure

Listener participants attended two sessions at the University of Canterbury Speech Perception Lab of between 60 and 90 minutes length. The second session took place within 1 to 21 days of the first session (mean = 5.28, SD = 4.89). Listeners received a $50 voucher on completion of the second test session. The first session included obtaining written consent and collecting demographic information, followed by half of the cognition and vocabulary tasks and half of the speech perception task. The second session involved the remaining half of the cognition and vocabulary tasks and the speech perception task. The order of the cognition and vocabulary tasks and the speech perception task was counterbalanced. The order of puretone audiometry was also counterbalanced between participants. Participants were assigned to one of eight test orders. Tasks were grouped to include similar tasks together and all of the WAIS-IV subtests were administered in the same order, as per WAIS-IV protocols (Wechsler, 2008).

Speech perception experiment. The speech perception experiment took place in a soundproof booth. Listeners were seated with a speaker at 0° azimuth at a distance of 0.5m. At the beginning of each session, the output of the sound system was measured with a 1000-Hz calibration tone and sound level meter (Reed ST-805 Compact Digital). The system gain was adjusted so that the speech level would be 65 dB A at the listener's head. Speech stimuli were presented to the speaker via an external soundcard (THX Trustudio PRO) and amplifier (Crown D-75A). A free field approach was used as the larger study included older adults,

---

[2] There were 4096 unique combinations of phrase (128), speaker (8), and SNR(4). 3479 of these combinations were presented in total. Each unique phrase was presented on average 1.84 times (SD = 0.9, range = 1 to 7).

providing a more comfortable environment for those unfamiliar with experiment participation.

Listeners were instructed that they would hear English phrases which were grammatically correct but did not make sense. Listeners were advised that some phrases would be easier to understand than others and encouraged to guess if they were unsure. They were instructed to repeat the phrase, and their spoken responses were recorded for later transcription with a Sony IC digital recorder. If participants were not sure about some of the words, they were asked to say the word “something” as a filler for each word that was unclear, as per previous studies (McAuliffe et al., 2013, 2014). Listeners were shown ten orthographic examples of the types of phrases they would hear, and played five example audio recordings of phrases mixed in noise at +6 dB SNR to familiarise listeners with the stimuli. Each phrase was presented once followed by 5s silence in which the listeners could respond.

During the first session, listeners heard the first half of their assigned list, two SNR blocks of 32 phrases. In the second session, listeners heard the remaining two SNR blocks of 32 phrases. After each block the listeners were able to take a short break before the next block began.

Data Analysis

Transcription. Each of the listener’s spoken responses was orthographically transcribed by the tester. Transcriptions were all lowercase; spelling was kept consistent with New Zealand English; and for all words or non-words whose pronunciation was not clear from the spelling, the orthographic transcription was accompanied by DISC phonetic transcription to ensure clarity for other researchers. Each instance of “something” was transcribed as “X”.

In order to reduce transcriber bias and improve transcription reliability, a second transcriber, blinded to the first transcription, also transcribed the responses. The two transcriptions were then automatically checked for identical matches. The second transcriber then checked any mismatches for spelling errors, fixed errors, and flagged the remaining mismatches for consensus checking. A third researcher completed the consensus checking. This involved listening to the mismatched responses and agreeing which of the two transcriptions was correct. The consensus checker produced the final transcription file for scoring and statistical analysis.

Dependent variable.

Word accuracy. The accuracy of individual content words in the phrases were scored automatically via a MATLAB (Mathworks, n.d.) script. The script checked whether the word in the transcription identically matched a word in the target phrase. Correct words received a score of 1, and incorrect words a score of 0.

Statistical analysis. A series of binomial mixed effects models using the lme4 package (Bates, Maechler, Bolker, & Walker, 2013) in R (version 2.15.3,(R Core Team, 2013)) were run to assess the effects of individual differences of vocabulary and cognition on word recognition and lexical cue use.

Mixed effects models were chosen so that phonological similarity and lexical frequency could be treated as continuous variables. This is in contrast to many previous studies (e.g., Bradlow & Pisoni, 1999; Dirks, Takayanagi, & Moshfegh, 2001; Dirks, Takayanagi, Moshfegh, et al., 2001; Freedman & Barlow, 2013; McArdle & Wilson, 2008; Sommers & Danielson, 1999; Takayanagi et al., 2002) which have utilised median splits to determine high and low lexical frequency and neighbourhood density groups. However, the use of these categorical variables may not be sensitive enough to determine the fine grained effects of phonological similarity and frequency on speech recognition in adverse listening conditions. In fact, one study (Bradlow & Pisoni, 1999) noted overlap between their two groups on all lexical characteristics: lexical frequency, neighbourhood density, and neighbourhood frequency. Furthermore, one word occurred in both the "high" and "low" groups, bringing into question the utility of the groups. Although there is one study which investigated neighbourhood density as a continuous variable using mixed-effects models (Janse & Newman, 2013), their results are limited to the use of non-words. Thus, the current study extends this work to English words and their PLD and lexical frequency measures.

Prior to conducting this analysis, correlations between the vocabulary and cognitive measures were checked to assess for multicollinearity in the data. Furthermore, the correlation between lexical characteristics of PLD and lexical frequency was examined in order to determine the extent to which they could be individual predictors. For the primary analyses, a binomial logit function was employed to account for the dependent variable of word accuracy (0 = incorrect, 1 = correct) (Jaeger, 2008). The models were fit with the maximum likelihood criterion. Composite scores of working memory, non-verbal intelligence, and processing speed were created for use as predictors in the model. A working memory composite score was calculated using the raw scores of the digit span (DS) and RSPAN tasks, e.g $\frac{Raw\ Score\ DS}{Total\ possible\ raw\ score\ DS} + \frac{Raw\ Score\ RSPAN}{Total\ possible\ raw\ score\ RSPAN}$. The remaining composite scores were calculated in the same way. The non-verbal intelligence measure was created from the raw scores of the block design and matrix reasoning tasks. The processing speed composite score was created from the raw scores of the symbol search and coding tasks. The individual measures of vocabulary knowledge, working memory, non-verbal intelligence, and processing speed were centred around their mean as per Neger et al. (2014).

Results

Vocabulary and Cognitive Measures Results

Listeners' average scores on measures of vocabulary and cognition are reported in Table 1. There was a range of scores across participants. Prior to conducting the primary analysis the relationships between the measures of vocabulary, working memory, non-verbal intelligence and processing speed were first examined to check for multicollinearity within the dataset. Intercorrelations between all measures are reported in Table 2. All intercorrelations were < 0.32, and no significant correlations were observed between the measures. This lack of intercorrelation allows for investigation of the separate effects of each of these measures on word recognition, and their ability to predict linguistic cue use.

Table 1. Mean scores for vocabulary and cognitive measures.

Measure	M	SD	Range
Vocabulary	112.96	10.28	88-139
Working Memory	1.36	0.15	1.00-1.69
Non-verbal intelligence	1.61	0.20	1.04-1.97
Processing Speed	1.23	0.18	0.89-1.79

Note: Vocabulary measured using the PPVT-IV (Dunn & Dunn, 2007).
Working memory is a composite score of the raw scores of the digit span (WAIS-IV) and RSPAN tasks.
Non-verbal intelligence is a composite score of the raw scores of the block design and matrix reasoning tasks (WAIS-IV).
Processing speed is a composite score of the raw scores of the symbol search and coding tasks (WAIS-IV).

Table 2. Pearson's correlation coefficients for vocabulary and cognitive measures.

Measure	Vocabulary	Working Memory	Non-verbal intelligence	Processing Speed
Vocabulary	1			
Working memory	0.14	1		
Non-verbal intelligence	0.31^{*}	0.20	1	
Processing Speed	0.21	0.11	0.22	1

Note: All correlations are non-significant when adjusted for multiple tests.
$^{*}p < 0.05$, non-adjusted for multiple tests.

Word Recognition Accuracy

Range of performance. Overall, listeners showed variation in their performance on the word recognition task (see Figure 1). Figure 1 demonstrates the proportion of words correctly identified by listeners averaged across the four SNR conditions (-5,-2, +1, and +4 dB SNRs**)**. Listeners' proportion of words correct ranged from 0.41 to 0.61 (M = 0.53, SD = 0.04).

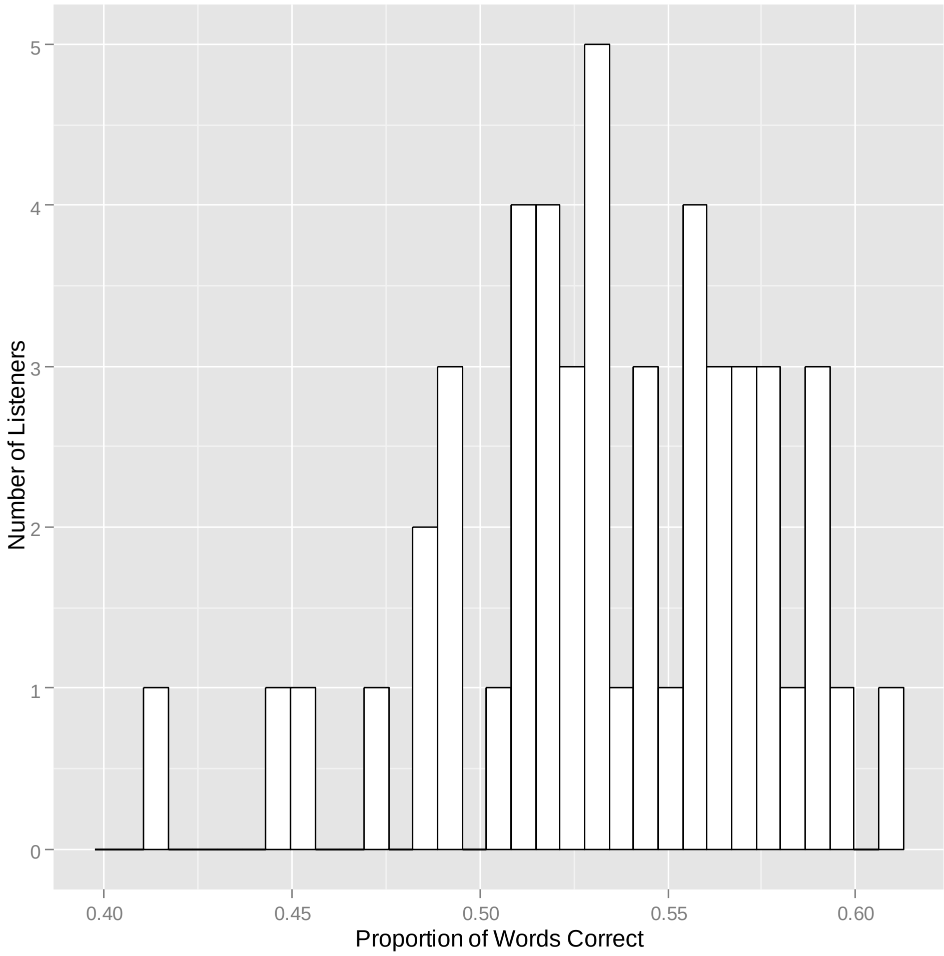

Figure 1. Histogram of the count of the proportion of words correctly recognised collapsed across four SNRs (-5,-2,+1,+4 dB SNRs).

Variation across SNR. As expected word recognition increased as the SNR became more favourable (see Figure 2). Figure 2 details listeners' proportion of words correctly identified across SNR blocks. These results confirm that the SNRs selected for the task provided a range of word recognition scores across listeners without the influence of floor and ceiling effects.

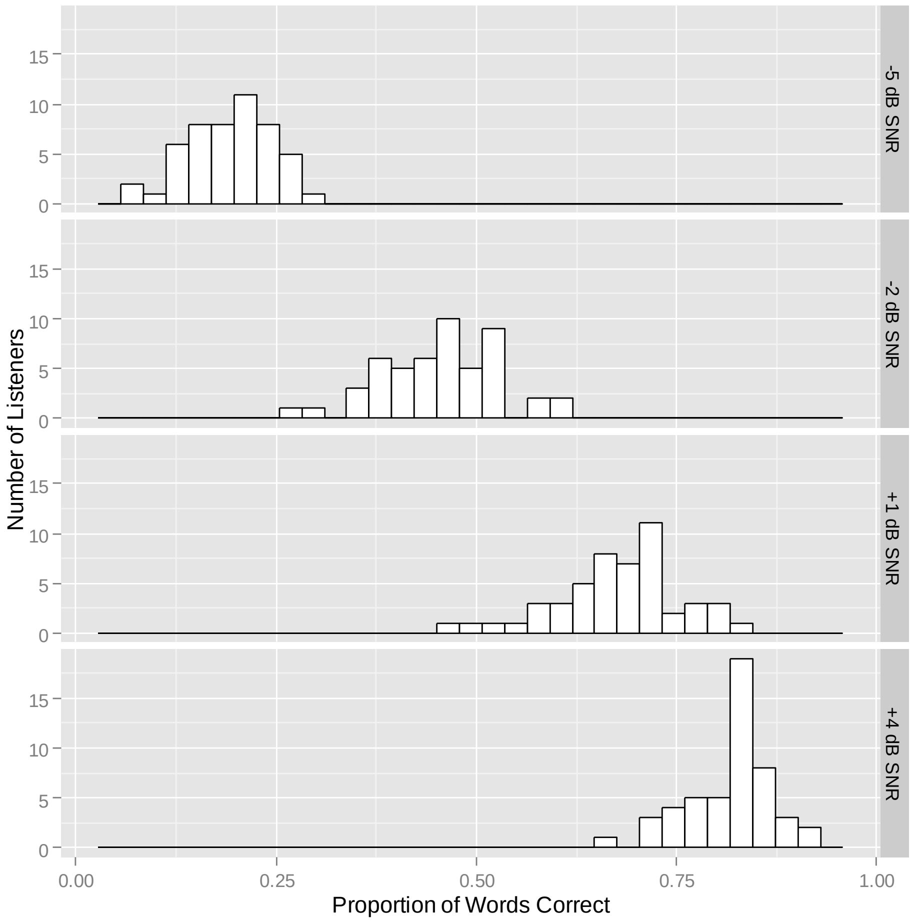

Figure 2. Histogram of the count of proportion of words correctly recognised by SNR.

Predictors of listener performance. The primary aim of the study was to determine whether listeners' individual differences with respect to vocabulary and cognition interacted with lexical cues to predict word recognition abilities. Binomial mixed effects modelling was used to address the hypotheses. The individual measures of vocabulary, the composite scores of working memory, non-verbal intelligence and processing speed, as well as the lexical characteristics (PLD and lexical frequency) and SNR, were entered simultaneously as fixed effects into a binomial mixed effects model, and all two-way interactions between fixed effects were tested. Random effects of participant, talker, phrase, and target word were included and random participant slopes of SNR were included to account for individuals responding differently across SNR blocks (Barr, Levy, Scheepers, & Tily, 2013; Cunnings, 2012). The model of best fit was determined through backward stepwise model selection. First, non-significant interactions were removed ($p > 0.05$), followed by single predictors. Model comparisons were used to confirm the increase in model fit at each change to the fixed effect structure.

Prior to conducting theses analyses, further checking was required. Since the effects of lexical frequency and PLD on word recognition were to be investigated, their correlation with one another was checked. As lexical frequency and PLD were only weakly correlated with one another (Pearson's $r = -0.10$, $p = 0.05$), they were entered as separate predictors into the model.

The final model for word recognition is detailed in Table 3. The estimates represent the effect size. Standard error, z-values and p-values are also reported. Given the uncertainty as to which method is best to report the significance of mixed effects models with maximal random slope structures, the reported p-values were derived from likelihood ratio tests between a model which included the specific fixed effect or interaction of interest and a model that does not contain the fixed effect or interaction but maintains all of the other model parameters, as per Neger et al. (2014). The final model revealed that as listening conditions became more favourable, listeners' word recognition improved ($p < 0.001$), complementing Figure 2. Individual listener characteristics of processing speed, working memory and non-verbal intelligence all demonstrated fixed main effects on word recognition. As processing speed and working memory increased, word recognition improved, ($p = 0.05$) and ($p < 0.05$) respectively. The effect of working memory was also moderated by lexical frequency; as working memory increased, higher word frequency was detrimental to word recognition accuracy ($p < 0.05$).

In contrast non-verbal intelligence demonstrated the opposite effect to working memory: higher non-verbal intelligence resulted in a decrease in word recognition accuracy ($p < 0.05$). However, non-verbal intelligence was moderated by lexical frequency, such that listeners with higher non-verbal intelligence exhibited better word recognition for higher frequency words ($p < 0.05$). There was no significant main effect for vocabulary knowledge on word recognition, although there was an interaction between vocabulary knowledge and SNR. Listeners with higher vocabularies showed increasing word recognition accuracy as SNRs became more favourable ($p < 0.05$). As already noted, word frequency's effect upon word recognition was moderated by the individual measures of working memory and non-verbal intelligence; however, there was also an interaction between word frequency and SNR. Higher frequency words were recognised more accurately as listening conditions became more favourable ($p < 0.05$). Furthermore, there was an overall main effect of word frequency: higher frequency words were recognised more accurately than lower frequency words ($p < 0.05$). Lastly, there were no interactions between any of the individual listener measures and PLD. Nevertheless, PLD demonstrated a significant main effect on word recognition accuracy, indicating that the more dissimilar a word was to other words, the more likely it was to be identified correctly ($p < 0.001$).

Table 3. Binomial mixed effects model for word recognition accuracy.

Fixed Effects	Estimate	SE	z	p
(Intercept)	-5.141	0.582	-8.830	< 0.001
Vocabulary	-0.006	0.006	-1.028	0.292
SNR	0.857	0.170	5.035	< 0.001
Lexical frequency	0.121	0.051	2.360	0.015
Phonological Levenshtein’s distance (PLD)	0.653	0.091	7.161	< 0.001
Processing speed	0.341	0.170	2.008	0.053
Working memory	2.126	0.898	2.369	0.019
Non-verbal intelligence	-1.435	0.686	-2.093	0.040
Vocabulary x SNR	0.005	0.002	2.231	0.028
Lexical frequency x SNR	0.035	0.016	2.195	0.036
Working memory x lexical frequency	-0.175	0.088	-1.986	0.049
Non-verbal intelligence x lexical frequency	0.160	0.067	2.383	0.020

Discussion

The main aim of this study was to examine the link between individual listener differences, and the use of the lexical cues of phonological similarity and word frequency, in speech recognition in adverse listening conditions for younger adults with normal hearing acuity. It was hypothesised that listeners' word identification would improve as listening conditions became more favourable. Furthermore, it was predicted that individual differences would vary with respect to their influence upon word recognition accuracy. In particular, increased vocabulary and working memory were expected to demonstrate increased word recognition accuracy; whereas processing speed and non-verbal intelligence were not expected to affect word recognition accuracy. The lexical characteristics of frequency and phonological similarity were anticipated to show different effects upon word recognition accuracy. Higher frequency words were predicted to be identified more readily than lower frequency words. On the other hand, words that were more phonologically similar to other words were expected to be identified with lower word accuracy. Finally, it was hypothesised that listener characteristics and lexical factors would interact to either aid or hinder word recognition accuracy. In particular, it was expected that better working memory would increase identification of higher frequency items; whereas poor working memory was expected to hinder the identification of items with increased phonological similarity to other words. On the other hand, the individual characteristics of non-verbal intelligence and processing speed were not predicted to interact with lexical characteristics. Given that the link between vocabulary and language-based factors was largely unexplored, the current study investigated the extent to which these interactions explained word recognition further.

The main findings of the study demonstrated that: (1) listeners' word identification improved as listening conditions became more favourable, and there was a range in word recognition abilities across listeners; individual listener characteristics had differing effects on word recognition with (2) increased processing speed and increased working memory both resulting in increased word recognition accuracy; whereas (3) increasing non-verbal intelligence was detrimental to word recognition accuracy; although (4) vocabulary knowledge alone had no main effect on word recognition, as the listening conditions became more favourable, higher vocabulary listeners identified more words correctly; with respect to lexical characteristics, (5) the more phonologically distinct a word was, the more likely it was to be correctly identified, and (6) the more often a word appeared in the language, the more likely it was to be recognised; (7) higher frequency words were also identified more accurately at higher SNR levels; however, these individual listener and language characteristics were not always working in isolation upon word recognition accuracy; (8) the positive effect of working memory on word recognition was reversed by increasing word frequency; on the other hand (9) non-verbal intelligence's negative influence on word recognition was reversed by increasing word frequency. These findings will be discussed in the following sections.

Word Identification in Adverse Conditions

The finding that word recognition improved as SNR increased was expected, and consistent with the vast body of psychoacoustic literature (e.g., Miller, 1947; Miller et al., 1951; Plomp, 1978). As the masking noise was reduced, there was less interference with the desired signal, and more of the signal was audible. Nevertheless, in spite of all listeners having puretone thresholds within normal limits, there was still a range of word recognition abilities overall, and within individual SNR blocks. Again, this finding was consistent with previous studies investigating the effects of listener characteristics on speech recognition. For example, in the current study the proportion of words correctly identified ranged from 0.41 to 0.61; in comparison, Tamati et al.'s (2013) poor and good listeners together reported a slightly wider range of 40% to 76% word recognition accuracy for their highly variable sentences. The reasons for this variation in listeners' word recognition abilities are discussed in the following sections.

Listener Characteristics' Influence on Word Recognition Accuracy

Listener characteristics were predictive of word recognition ability, although this was not always in agreement with the literature. Each of the individual characteristics of working, memory, processing speed, non-verbal intelligence and vocabulary knowledge are discussed below.

Working memory. The current study found that as working memory increased, word recognition accuracy in noise improved. This finding is consistent with other research (e.g., Ellis & Munro, 2013; Tamati et al., 2013), and in line with the presented hypothesis. However, these results contrast those found by another similar study (McAuliffe et al., 2013). McAuliffe et al. (2013) also employed non-sensical phrases[3], yet found no influence of working memory, as measured by the Wechsler Memory Scale digit span and letter-number sequencing tasks (Wechsler, 1997b). The authors suggested that the short phrases and reduced variation in listeners working memory scores may have accounted for the lack of a finding. In the current study, however, the working memory measure included both a digit span component and reading span task. The reading span task has been shown to correlate with speech perception in adverse conditions (Akeroyd, 2008). Therefore, it is possible that the reading span component of the working memory composite score was better able to capture the differences in word recognition ability than the digit span measure alone.

The current study also differed from McAuliffe et al. (2013) with respect to the number of participants and type of stimuli degradation utilised. The current study included a larger number of participants (n = 50) compared to McAuliffe et al. (2013) (n = 16 younger, n

[3] The current study employed 47 of the 60 phrases used by McAuliffe et al. (2013).

= 16 older listeners analysed in separate models). This increased power of the current study could account for the difference in findings. Furthermore, the current study employed systematically degraded stimuli by mixing a talker's speech with their long term average speech spectra. In contrast, McAuliffe et al.'s (2013) study included naturally degraded speech stimuli, phrases spoken by participants with hypokinetic dysarthria. Although including naturally degraded phrases adds to the ecological validity of studies, it also introduces issues. The hallmarks of hypokinetic dysarthria are its unnaturally fast rate of speech, reduced pitch variation, reduced loudness and imprecise phoneme articulation (Darley, Aronson, & Brown, 1969). Thus, perhaps this natural degradation provided too much of an unfamiliar listening environment and listeners were unable to utilise their working memory efficiently. For example, even if listeners were able to hold the speech stream in working memory long enough to use, perhaps they did not have enough suitable candidates in long-term memory to match the target.

Another study investigating word recognition, Krull et al. (2013), also did not find a link between working memory, as measured by the working memory test battery (Lewandowsky, Oberauer, Yang, & Ecker, 2010), which also includes a reading span task; however, the authors note that their lack of a relationship is consistent with studies which have used isolated words as stimuli, as summarised by Akeroyd (2008). So although the current study investigated word accuracy the fact that the words were embedded in non-sensical phrases may have strengthened the link between word recognition and working memory.

The current findings could be paired with results from a study by Francis (2010) who employed a dual-task paradigm investigating the effect of increasing working memory load had on listeners' abilities to successfully recognise speech in the presence of a competing talker. The researchers found that increasing cognitive load lead to increased interference from the distracting speech signal. These findings could suggest that one reason that listeners with increased working memory performed better in the present study is because they were better able to limit noise as a distraction. This could be because listeners with increased working memory were better able to hold the portions of glimpsed speech in memory for longer in order to successfully piece them together.

Processing speed. Increased processing speed was also found to be beneficial to word recognition. This finding is in contrast to the present hypothesis and in conflict with a number of studies who report no effect of processing speed on speech recognition for younger listeners (Benard et al., 2014; Neger et al., 2014; Sommers & Danielson, 1999). However, the current study differs to previous studies in a number of ways which could account for the different findings. For example, in the current study there was no correlation between processing speed and the other individual measures of vocabulary and cognition, whereas in Neger et al.'s (2014) study processing speed was significantly correlated with working memory, and neither of these were significant predictors in that study. Furthermore, the use of low context phrases in the current study, could account for differences to previous studies: Benard et al. (2014) utilised high context sentences, as did Neger et al. (2014) for noise-vocoded stimuli. In contrast, Sommers and Danielson (1999) utilised both high and low predictability sentences. However, their measure of processing speed was based on the response latencies of participants completing the recognition tasks, and not upon a separate task. One study (Desjardins & Doherty, 2013) to report processing speed effects on younger listeners' speech recognition in noise utilised similar tasks to the current study, the digit symbol substitution test from the WAIS-III (Wechsler, 1997a) and both high and low context sentences (R-SPIN) (Bilger, Nuetzel, Rabinowitz, & Rzeczkowski, 1984). Furthermore, although the current study's finding that younger listeners with higher processing speed demonstrated increased speech recognition ability is inconsistent with much of the literature, the effect is in a direction which lends itself to a possible explanation. Listeners with increased processing speed may be able to utilise glimpses in the speech signal much more readily than listeners with lower processing speeds. Listeners only had 5 seconds in which to respond before the next phrase was automatically presented, and each phrase was approximately 2.37s in duration. This limited time may suggest listeners with increased processing speed were better able to piece together the glimpses of speech to form lexical items.

The current findings are also consistent with previous studies which report the benefits of increased processing speed upon speech recognition in older adults (e.g., Desjardins & Doherty, 2013; Helfer & Freyman, 2014) and listeners with a wide range of ages (e.g., Woods et al., 2013). Helfer and Freyman (2014) utilised the Connections Test (Salthouse et al., 2000) and found a link between processing speed and word recognition for their older listeners; however, they did not compare processing speed with word recognition scores for their younger listeners who had significantly better processing scores than the older listeners. In addition to their finding that increased processing speed aided younger listeners' speech recognition in noise, Desjardins and Doherty (2013) also found that increased processing speed was associated with older listeners' improved speech recognition in noise performance. Increased processing speed also negatively correlated with listening effort (Desjardins & Doherty, 2013), i.e., listeners with increased processing speed found the listening tasks less effortful. Furthermore, a recent longitudinal study (Pronk et al., 2013) following older adults for 3 to 7 years found that declining processing speed accounted for a large proportion of the changes to listeners' speech recognition in noise abilities over time. Given that processing speed appears to account for more differences in older listeners it is of interest to extend the current study to include older listeners, in order to determine if the effect of processing speed is stronger for these low predictability stimuli.

Non-verbal intelligence. However, not all listener characteristics in the current study had a positive influence on word recognition accuracy. Listeners with increased non-verbal intelligence were more likely to incorrectly identify words. Like the findings for processing speed, the non-verbal intelligence results are in contrast to those found in the speech recognition literature (e.g., Conway et al., 2010; Tamati et al., 2013) who found no link between non-verbal intelligence and performance. The findings are also different to the current study's prediction that non-verbal intelligence would not influence speech recognition. However, unlike processing speed, the direction of the effect of non-verbal intelligence does not readily lend itself to an explanation. In spite of this, non-verbal intelligence does interact with lexical frequency, which will be discussed further below.

Vocabulary. In contrast to predictions, the current study failed to find a significant main effect of receptive vocabulary upon speech recognition accuracy. This finding is in contrast to a number of studies who have found positive effects of receptive vocabulary overall, (e.g., Alamsaputra et al., 2006; McAuliffe et al., 2013). However, the current study did find that as the SNR became more favourable, higher vocabulary knowledge aided word recognition. One reason for this finding could be that in the lower SNRs, perhaps there was not enough information for listeners to glimpse to even allow vocabulary knowledge to be effective; at the lowest SNR, listeners only correctly identified on average 0.19 of the words presented. Mattys, White, and Melhorn (2005) suggested that as listening conditions improve, listeners change their reliance on using sublexical knowledge (e.g., acoustic-phonetic cues) to lexical knowledge, which could explain the influence of vocabulary knowledge on speech recognition at higher SNR levels.

Language Characteristics' Influence on Word Recognition Accuracy

Both lexical cues of word frequency and PLD had clear effects on listeners' word recognition accuracy. One possible reason for this is that the current study employed the use of mixed-effects models which allows for a more detailed analysis of listener and language characteristics, rather than splitting lexical frequency and phonological similarity into strictly high and low groups commonly utilised in previous studies (e.g., Dirks, Takayanagi, & Moshfegh, 2001; Dirks, Takayanagi, Moshfegh, et al., 2001; Freedman & Barlow, 2013; McArdle & Wilson, 2008; Sommers & Danielson, 1999; Takayanagi et al., 2002). By treating these measures as continuous variables, the current study may have found subtler influences of listener characteristics and language-based factors on word recognition accuracy.

Lexical frequency. Higher word frequency aided word recognition overall, which is consistent with previous studies (e.g., Goldinger et al., 1989; Howes, 1957; Taler et al., 2010), and in line with this study's hypothesis. Furthermore as the SNR increased, higher frequency words became even more recognisable than lower frequency words. Whereas Taler et al. (2010) found that the difference in accurate identification between low frequency items and high frequency items was more prevalent at a lower SNR (-3 dB). Nevertheless, the current finding is in keeping with Mattys et al.'s (2005) hierarchy in that lexical cues are employed when listening conditions become favourable enough for them to dominate. Word frequency also interacts with the individual factors of working memory and non-verbal intelligence, which will be discussed in the section below.

Phonological similarity. The finding that increasing phonological dissimilarity aids word recognition is consistent with the only other study to investigate PLD in relation to spoken word recognition (Suárez et al., 2011) and in keeping with the current study's prediction. This finding also complements the extensive literature on neighbourhood density, which state that words with fewer neighbours are easier to understand (e.g., Luce & Pisoni, 1998; Vitevitch & Luce, 1998, 1999). The current finding is logical, in that listeners have fewer choices to make when selecting words that have fewer phonological competitors. The result is less competition from similar sounding words in a listener's lexicon, leading to more accurate word identification. It also further suggests that PLD is a suitable measure of phonological similarity.

Given the findings of the interactions of the individual predictors of vocabulary and lexical frequency with SNR, it may have been expected to find an interaction of PLD and SNR. Perhaps the lack of an interaction for PLD with SNR was due to limited variance in the stimuli PLDs, as suggested by McArdle and Wilson (2008). Two studies which have found interactions between phonological similarity and SNR (Krull, Choi, Kirk, Prusick, & French, 2010; Taler et al., 2010) found that these effects were larger at lower SNRs.

The Link between Listener Characteristics and Language-based Factors' Influence on Word Recognition Accuracy

Only a small number of studies have investigated the influence of listener characteristics and lexical cue use together in speech recognition (Bradlow & Pisoni, 1999; Frisch et al., 2001; Janse & Newman, 2013; McAuliffe et al., 2013; Sommers & Danielson, 1999; Taler et al., 2010). Hence, the present study was largely exploratory regarding predictions of the interactions between listener and lexical characteristics.

In the current study, the positive effect of increasing lexical frequency on word recognition was moderated by the listener characteristics of working memory and non-verbal intelligence. However, these listener characteristics had differing influences on word recognition accuracy. As working memory increased, higher frequency words were more likely to be incorrectly identified. In contrast to predictions, this result differs from Taler et al. (2010) who found increased short term memory, as measured by a forward digit span task, resulted in better identification of high frequency words at a higher SNR. However, the working memory measure in the present study incorporated a verbal working memory component (RSPAN) which could possibly account for the different finding. Another possible reason for the current finding could be that listeners with higher working memory utilise this working memory capacity to hold a larger number of lexical candidates from their lexicon. The issue arises in that these possible candidates are likely to be higher frequency words, thus creating more competition. Thus, although, high frequency words and increased working memory are beneficial for word recognition when isolated, together they work against word recognition. It should also be noted that this effect is only just significant ($p = 0.049$), thus with more participants and increased statistical power this could change.

In spite of this interaction between working memory and word frequency, there was no interaction between working memory and PLD as was hypothesised. However, in line with predictions, there was no interaction with processing speed and lexical cues.

Also in contrast to predictions, non-verbal intelligence did interact with lexical characteristics to influence word recognition accuracy. As non-verbal intelligence increased, higher frequency items were identified more accurately. Therefore, although non-verbal intelligence in isolation was detrimental to word recognition accuracy, increasing non-verbal intelligence appears to combat the negative influence that working memory has upon word recognition for higher frequency items. This suggests that listeners may call on different areas of cognition to aid in different listening situations depending on which lexical cues are available or provide most benefit at the time.

The current study also failed to find a link between vocabulary knowledge and lexical cue use. This was despite studies (e.g., Bradlow & Pisoni, 1999; Neger et al., 2014) whose work had suggested the possibility of a link. Bradlow and Pisoni (1999)'s findings that hard word recognition accuracy depended upon the length of time spent in an English speaking environment, whereas a listeners' familiarity of hard words depended upon the age at which a listener began to formally study English, suggested that hard word familiarity and hard word recognition were influencing different areas of language proficiency. Their findings taken together with the findings of Neger et al. (2014) that vocabulary and statistical learning individually predicted listeners' perceptual learning suggested that higher vocabulary listeners may use certain other underlying cues in language to aid their speech perception in adverse conditions. However, given the current study failed to find a link between vocabulary knowledge and lexical cue use, the underlying mechanisms between vocabulary and speech recognition are still unclear.

Although the current study failed to find links between vocabulary knowledge and lexical cue use, there were links found between lexical frequency and both working memory and non-verbal intelligence. These findings take a step closer toward understanding the underlying mechanisms of individual differences and speech recognition.

Limitations of the Study and Future Research Directions

The current study's findings should be considered with respect to its limitations. Firstly, generalisations are limited to young listeners with normal hearing. In order to consider the effects of ageing and hearing loss, the current study should be extended to include older adults and listeners with hearing impairment. Older adults are more likely to show larger deficits in cognitive abilities than younger adults (e.g., Neger et al., 2014) and thus demonstrate greater effects of individual differences upon word recognition. Given that the effects of cognition on speech recognition are often secondary to hearing loss (e.g., Akeroyd, 2008), investigating a population with hearing impairment would further add to the picture of the role of cognition and speech recognition in adverse listening conditions.

Secondly, due to limitations in the length of test sessions, there was no measure of attention-switching, or inhibition included in the present study. These measures may have provided further information about the ability of listeners to utilise lexical cues, as they have been found to relate with phonological similarity in the past (Janse & Newman, 2013; Sommers & Danielson, 1999).

Although a focus of the present study was to investigate lexical cue use for items that were representative of everyday language, future studies could utilise stimuli more specifically chosen for their lexical frequency and phonological similarity differences.

Another possibility is to investigate the link between individual differences and the lexical properties of listeners' misperceptions in order to further understand listener speech recognition strategies. This would build upon the findings of Vitevitch (2002) who found that when listeners incorrectly identified words they responded with items that although did not differ in neighbourhood density and frequency properties compared to the target words, were higher in frequency and had higher neighbourhood density than other words in the language.

The present study could also be extended to include additional adverse listening conditions, such as speech in babble or competing talker in order to assess the effects of informational masking on individual differences and lexical cue use.

Finally, despite having hearing thresholds within normal limits for puretone audiometry, many normal hearing listeners report listening difficulties in noisy environments. Thus, it would be worthwhile investigating self-perceived hearing ability of these listeners.

Conclusion

The current study highlighted a wide range of word recognition abilities in this younger listener population. The findings demonstrated that word recognition is a complex process influenced by language and listener-based factors simultaneously. Furthermore, the interactions between lexical frequency and both working memory and non-verbal intelligence, provide further insight into individual listener differences and lexical cue use in word recognition.

The current study may have found subtler influences of listener characteristics and language-based factors on word recognition accuracy due to its design. This suggests that by investigating a population of older listeners and listeners with hearing impairment, the results would be even more varied, and provide further knowledge of the underlying mechanisms of individual differences and speech recognition. Further understanding of how these listener differences affect an individual's speech processing may lead to the development of improved signal processing techniques and rehabilitation strategies.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, *47*(S2), S53–S71. doi:10.1080/14992020802301142

Alamsaputra, D. M., Kohnert, K. J., Munson, B., & Reichle, J. (2006). Synthesized speech intelligibility among native speakers and non-native speakers of English. *Augmentative and Alternative Communication*, *22*(4), 258–268. doi:10.1080/00498250600718555

Assmann, P., & Summerfield, Q. (2004). The perception of speech under adverse conditions. In *Speech processing in the auditory system* (pp. 231–308). New York, NY: Springer.

Auer, E. T., & Luce, P. A. (2005). Probabilistic phonotactics in spoken word recognition. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 610–630). Blackwell Publishing Ltd. Retrieved from http://onlinelibrary.wiley.com

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database [CD-ROM]*. Philidelphia: Linguistic Data Consortium.

Baddeley, A. (1998). Working memory. *Comptes Rendus de l'Académie Des Sciences - Series III - Sciences de La Vie*, *321*(2–3), 167–173. doi:10.1016/S0764-4469(97)89817-4

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340–357. doi:10.1037/0096-1523.10.3.340

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English lexicon project. *Behavior Research Methods*, *39*(3), 445–459.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). *lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-4*. Retrieved from http://cran.r-project.org/web/packages/lme4/index.html

Benard, M. R., Mensink, J. S., & Baskent, D. (2014). Individual differences in top-down restoration of interrupted speech: Links to linguistic and cognitive abilities. *The Journal of the Acoustical Society of America*, *135*(2), EL88–EL94. doi:10.1121/1.4862879]

Benichov, J., Cox, L. C., Tun, P. A., & Wingfield, A. (2012). Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability and cognitive function. *Ear and Hearing*, *33*(2), 250–256. doi:10.1097/AUD.0b013e31822f680f

Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzeczkowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech, Language, and Hearing Research*, *27*(1), 32–48. doi:10.1044/jshr.2701.32

Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *The Journal of the Acoustical Society of America*, *106*(4), 2074–2085. doi:10.1121/1.427952

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, *109*(3), 1101–1109. doi:10.1121/1.1345696

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. doi:10.3758/BRM.41.4.977

CHABA (Committee on Hearing and Bioacoustics, Working Group on Speech Understanding and Aging). (1988). Speech understanding and aging. *J Acoust Soc Am*, *83*, 859–895. doi:10.1121/1.395965

Collison, E. A., Munson, B., & Carney, A. E. (2004). Relations among linguistic and cognitive skills and spoken word recognition in adults with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *47*(3), 496–508. doi:10.1044/1092-4388(2004/039)

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*(3), 356–371. doi:10.1016/j.cognition.2009.10.009

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, *28*(3), 369–382. doi:10.1177/0267658312443651

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. doi:10.1016/S0022-5371(80)90312-6

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech Language and Hearing Research*, *12*(2), 246–269. doi:10.1044/jshr.1202.246

Desjardins, J. L., & Doherty, K. A. (2013). Age-Related Changes in Listening Effort for Various Types of Masker Noises: *Ear and Hearing*, *34*(3), 261–272. doi:10.1097/AUD.0b013e31826d0ba4

Dirks, D. D., Takayanagi, S., & Moshfegh, A. (2001). Effects of lexical factors on word recognition among normal-hearing and hearing-impaired listeners. *Journal of the American Academy of Audiology*, *12*(5), 233–244.

Dirks, D. D., Takayanagi, S., Moshfegh, A., Noffsinger, P. D., & Fausti, S. A. (2001). Examination of the neighborhood activation theory in normal and hearing-impaired listeners. *Ear and Hearing*, *22*(1), 1–13.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (Fourth). San Antonio, TX: Pearson.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (Third). Circle Pines, MN: American Guidance Service.

Ellis, R. J., & Munro, K. J. (2013). Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition? *International Journal of Audiology*, *52*(1), 14–22. doi:10.3109/14992027.2012.721013

Francis, A. L. (2010). Improved segregation of simultaneous talkers differentially affects perceptual and cognitive capacity demands for recognizing speech in competing speech. *Attention, Perception, & Psychophysics*, *72*(2), 501–516. doi:10.3758/APP.72.2.501

Freedman, S. E., & Barlow, J. A. (2013). Effects of neighborhood density on adult word repetition. *International Journal of English Linguistics*, *3*(2), 1–13. doi:10.5539/ijel.v3n2p1

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language*, *51*(1), 136–158. doi:10.1016/j.jml.2004.03.008

Frisch, S. A., Large, N. R., Zawaydeh, B., & Pisoni, D. B. (2001). Emergent phonotactic generalizations in English and Arabic. In J. L. Bybee & P. J. Hopper (Eds.), *Typological studies in language* (Vol. 45, pp. 159–180). Amsterdam: John Benjamins Publishing.

Gilbert, J. L., Tamati, T. N., & Pisoni, D. B. (2013). Development, Reliability and Validity of PRESTO: A New High-Variability Sentence Recognition Test. *Journal of the American Academy of Audiology*, *24*(1), 26–36. doi:10.3766/jaaa.24.1.4

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, *28*(5), 501–518. doi:10.1016/0749-596X(89)90009-0

Helfer, K. S., & Freyman, R. L. (2014). Stimulus and listener factors affecting age-related changes in competing speech perception. *The Journal of the Acoustical Society of America*, *136*(2), 748–759. doi:10.1121/1.4887463

Howes, D. (1957). On the relation between the intelligibility and frequency of occurrence of English words. *The Journal of the Acoustical Society of America*, *29*(2), 296–305. doi:10.1121/1.1908862

Humes, L. E., Kidd, G. R., & Lentz, J. J. (2013). Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults. *Frontiers in Systems Neuroscience*, *7*(55), 1–16. doi:10.3389/fnsys.2013.00055

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. doi:10.1016/j.jml.2007.11.007

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *The Quarterly Journal of Experimental Psychology*, *65*(8), 1563–1585. doi:10.1080/17470218.2012.658822

Janse, E., & Newman, R. S. (2013). Identifying nonwords: Effects of lexical neighborhoods, phonotactic probability, and listener characteristics. *Language and Speech*, *56*(4), 421–441. doi:10.1177/0023830912447914

Jarvis, B. G. (2010). DirectRT (Version 2010.3.109). New York, NY: Empirisoft Corporation.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. doi:10.1037/0278-7393.20.4.824

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321–340. doi:10.1016/j.cognition.2010.05.011

Krull, V., Choi, S., Kirk, K. I., Prusick, L., & French, B. (2010). Lexical effects on spoken-word recognition in children with normal hearing. *Ear and Hearing*, *31*(1), 102–114. doi:10.1097/AUD.0b013e3181b7892f

Krull, V., Humes, L. E., & Kidd, G. R. (2013). Reconstructing wholes from parts: Effects of modality, age, and hearing loss on word recognition. *Ear and Hearing*, *34*(2), e14–e23. doi:10.1097/AUD.0b013e31826d0c27

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lewandowsky, S., Oberauer, K., Yang, L.-X., & Ecker, U. K. H. (2010). A working memory test battery for MATLAB. *Behavior Research Methods*, *42*(2), 571–585. doi:10.3758/BRM.42.2.571

Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. (1991). Word familiarity and lexical fluency: Individual differences in serial recall of spoken words. In *Research on Speech Perception Progress Rep. No. 17* (pp. 229–239). Bloomington: Indiana University, Psychology Department, Speech Research Laboratory.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1–36.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203–208. doi:10.3758/BF03204766

Mathworks. (n.d.). *MATLAB*. Natick, MA.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7-8), 953–978. doi:10.1080/01690965.2012.705006

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: a hierarchical framework. *Journal of Experimental Psychology: General*, *134*(4), 477–500. doi:10.1037/0096-3445.134.4.477

McArdle, R., & Wilson, R. H. (2008). Predicting word-recognition performance in noise by young listeners with normal hearing using acoustic, phonetic, and lexical variables. *Journal of the American Academy of Audiology*, *19*(6), 507–518. doi:10.3766/jaaa.19.6.6

McAuliffe, M. J., Gibson, E. M., Kerr, S. E., Anderson, T., & LaShell, P. J. (2013). Vocabulary influences older and younger listeners' processing of dysarthric speech.

The Journal of the Acoustical Society of America, *134*(2), 1358–1368. doi:10.1121/1.4812764

McAuliffe, M. J., Kerr, S. E., Gibson, E. M. R., Anderson, T., & Lashell, P. J. (2014). Cognitive-perceptual examination of remediation approaches to hypokinetic dysarthria. *Journal of Speech, Language, and Hearing Research: JSLHR*, *57*, 1268–1283. doi:10.1044/2014_JSLHR-S-12-0349

McAuliffe, M. J., & Sinex, D. G. (2014, 2017). Understanding how listeners comprehend distorted speech. Royal Society of New Zealand Marsden Grant.

Miller, G. A. (1947). The masking of speech. *Psychological Bulletin*, *44*(2), 105–129. doi:10.1037/h0055960

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, *41*(5), 329–335.

Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, *22*(2), 167–173. doi:10.1121/1.1906584

Misyak, J. B., & Christiansen, M. H. (2012). Statistical Learning and Language: An Individual Differences Study. *Language Learning*, *62*(1), 302–331. doi:10.1111/j.1467-9922.2010.00626.x

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, *118*(1), 43–71. doi:10.1037/0096-3445.118.1.43

Munson, B. (2001). Relationships between vocabulary size and spoken word recognition in children aged 3 to 7. *Contemporary Issues in Communication Science and Disorders*, *28*(20-29).

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. doi:10.1111/j.1532-5415.2005.53221.x

Neger, T. M., Rietveld, T., & Janse, E. (2014). Relationship between perceptual learning in speech and statistical learning in younger and older adults. *Frontiers in Human Neuroscience*, *8*(628), 1–17. doi:10.3389/fnhum.2014.00628

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. In *Research on Speech Perception Progress Report No. 10* (pp. 357–376). Bloomington: Indiana University, Psychology Department, Speech Research Laboratory.

Pichora-Fuller, M. K., & Singh, G. (2006). Effects of age on auditory and cognitive processing: implications for hearing aid fitting and audiologic rehabilitation. *Trends in Amplification*, *10*(1), 29–59. doi:10.1177/108471380601000103

Pisoni, D. B. (2007). *WordFam: Rating Word Familiarity in English*. Bloomington: Indiana University.

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America*, *63*(2), 533–549. doi:10.1121/1.381753

Pronk, M., Deeg, D. J. H., Festen, J. M., Twisk, J. W., Smits, C., Comijs, H. C., & Kramer, S. E. (2013). Decline in Older Persons' Ability to Recognize Speech in Noise: The Influence of Demographic, Health-Related, Environmental, and Cognitive Factors. *Ear and Hearing*, *34*(6), 722–732. doi:10.1097/AUD.0b013e3182994eee

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Saffran, J. R. (2003). Statistical language learning mechanisms and constraints. *Current Directions in Psychological Science*, *12*(4), 110–114. doi:10.1111/1467-8721.01243

Salthouse, T. A., Toth, J., Daniels, K., Parks, C., Pak, R., Wolbrette, M., & Hocking, K. J. (2000). Effects of aging on efficiency of task switching in a variant of the Trail Making Test. *Neuropsychology*, *14*(1), 102–111. doi:10.1037/0894-4105.14.1.102

Schlauch, R., & Nelson, P. (2009). Auditory brainstem response: Differential diagnosis. In J. Katz, L. Medwetsky, R. Burkard, & L. J. Hood (Eds.), *Handbook of clinical audiology* (6th ed., pp. 30–49). Baltimore., Maryland, USA: Lippincott Williams & Wilkins.

Schoof, T., & Rosen, S. (2014). The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. *Frontiers in Aging Neuroscience*, *6*(307), 1–14. doi:10.3389/fnagi.2014.00307

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *The Journal of the Acoustical Society of America*, *123*, 489–499. doi:10.1121/1.2783762

Sinex, D. G. (2013). Recognition of speech in noise after application of time-frequency masks: Dependence on frequency and threshold parameters. *The Journal of the Acoustical Society of America*, *133*(4), 2390–2396. doi:10.1121/1.4792143

Sommers, M. S., & Danielson, S. M. (1999). Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context. *Psychology and Aging*, *14*(3), 458–472. doi:10.1037/0882-7974.14.3.458

Spearman, C. (1904). "General Intelligence," Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201–292. doi:10.2307/1412107

Storkel, H. L. (2013). A corpus of consonant–vowel–consonant real words and nonwords: Comparison of phonotactic probability, neighborhood density, and consonant age of acquisition. *Behavior Research Methods*, *45*(4), 1159–1167. doi:10.3758/s13428-012-0309-7

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*, *49*(6), 1175–1192. doi:10.1044/1092-4388(2006/085).

Suárez, L., Tan, S. H., Yap, M. J., & Goh, W. D. (2011). Observing neighborhood effects without neighbors. *Psychonomic Bulletin & Review*, *18*(3), 605–611. doi:10.3758/s13423-011-0078-9

Takayanagi, S., Dirks, D. D., & Moshfegh, A. (2002). Lexical and talker effects on word recognition among native and non-native listeners with normal and impaired hearing. *Journal of Speech, Language & Hearing Research*, *45*(3), 585–597.

Taler, V., Aaron, G. P., Steinmetz, L. G., & Pisoni, D. B. (2010). Lexical neighborhood density effects on spoken word recognition and production in healthy aging. *Journals of Gerontology*, *65B*(5), 551–560. doi:10.1093/geronb/gbq039

Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2013). Some factors underlying individual differences in speech recognition on PRESTO: A first report. *Journal of the American Academy of Audiology*, *24*(7), 616–634. doi:10.3766/jaaa.24.7.10

Tamati, T. N., & Pisoni, D. B. (2014). Non-native listeners' recognition of high-variability speech using PRESTO. *Journal of the American Academy of Audiology*, *25*(9), 869–892. doi:10.3766/jaaa.25.9.9

Tompkins, C. A., Bloise, C. G., Timko, M. L., & Baumgaertner, A. (1994). Working memory and inference revision in braindamaged and normally aging adults. *Journal of Speech and Hearing Research, 37*(4), 896–912. doi:10.1044/jshr.3704.896

Unisyn lexicon. (n.d.). Centre for Speech Technology ResearchCentre for Speech Technology Research, University of Edinburgh. Retrieved from www.cstr.ed.ac.uk/projects/unisyn/

Vitevitch, M. S. (2002). Naturalistic and experimental analyses of word frequency and neighborhood density effects in slips of the ear. *Language and Speech*, *45*(4), 407–434. doi:10.1177/00238309020450040501

Vitevitch, M. S. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language, and Hearing Research*, *51*(2), 408–422. doi:10.1044/1092-4388(2008/030)

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*(4), 325–329. doi:10.1111/1467-9280.00064

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*(3), 374–408. doi:10.1006/jmla.1998.2618

Watson, B. U. (1991). Some relationships between intelligence and auditory discrimination. *Journal of Speech Language and Hearing Research*, *34*(3), 621–627. doi:10.1044/jshr.3403.621

Wechsler, D. (1997a). *WAIS-III, Wechsler Adult Intelligence Scale: Administration and Scoring Manual*. Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale (WMS-III)*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2008). *Wechsler adult intelligence scale–Fourth Edition (WAIS–IV)*. San Antonio, TX: NCS Pearson.

Woods, W. S., Kalluri, S., Pentony, S., & Nooraei, N. (2013). Predicting the effect of hearing loss and audibility on amplified speech reception in a multi-talker listening scenario. *The Journal of the Acoustical Society of America*, *133*(6), 4268–4278. doi:10.1121/1.4803859

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*(4), 502–529. doi:10.1016/j.jml.2009.02.001

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: a new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979. doi:10.3758/PBR.15.5.971

Appendix: Phrasal stimuli list

a reason guests contained

absent fields did morning

account for who could knock

address her meeting time

admit the gear beyond

advance but sat appeal

afraid beneath demand

age of centered wagons

allow assured remains

amend estate approach

ancient leading students

and spoke behind her sin

appear to wait then turn

around without such roads

assume to catch control

attack became concerned

attend the trend success

avoid or beat command

award his drain away

balance clamp and bottle

before his wish was strong

begun his crown belief

beside a sunken bat

birth the notice symbol

bolder ground from justice

bush is chosen after

butcher shook the middle

career despite research

cheap console in paper

closer showing metal

commit such used advice

compare events of bank

confused but roared again

connect the beer device

constant willing walker

cool the jar in private

create her spot of art	form object with knowing
darker painted baskets	fort believed such borders
debate reply was mean	frame her seed to answer
define respect instead	friendly moon was sectioned
degree prevents from games	functions aim his acid
depend is longer sound	had eaten junk and train
derived extent with streets	had value plants to mind
direct with sweet extreme	headed wheels with stories
distant leaking basement	her owners arm the phone
divide across retreat	higher patient concept
done with finest handle	hold a page of fortune
during pattern programs	housing drawn in samples
each informed from flowers	hundred printed license
effect his wage but stood	ideal conduct had songs
embark or take her sheet	improve in driving cloud
exam of joy began	indeed a tax ascent
extend but please his stones	its harmful note abounds
force of focus moment	kick a tab above them
forget the joke below	lake is pressure sofa

legal stain for distance

plenty peas or causes

mark a single ladder

pooling pill or cattle

mate denotes a judgement

present relief among

may the same pursue it

push her equal culture

mistake delight for heat

question major nature

mister type is fashion

rampant boasting captain

mode campaign for budget

refer to good from league

model sad and local

refused percent to goal

motion double garden

release between such trees

narrow seated member

remove and name for stake

obtain contracts from tasks

resting older earring

or spent sincere aside

rhythm under artist

orders fairly level

rocking modern poster

pain can follow agents

rode the lamp for testing

passing plus a factor

round and bad for carpet

perceive sustained supplies

rowing farther matters

permits achieved but lied

seat for locking runners

pick a chain for action

secure but lease apart

plan reduced its setting

sight about the cannon

sparkle enters broken

speaking clear is power

submit his cash report

suggest its price reserve

support with dock and cheer

target keeping season

technique but green result

tension known from pleasure

thinking charged the hearing

to sort but fear inside

transcend almost betrayed

unique exchange in holes

unless escape can learn

unseen machines agree

useful music riding

world repeats with feelings